Uniquely minimizing costs for the Kantorovitch problem *

A. Moameni[†] L. Rifford[‡]

Abstract

The purpose of the present paper is to establish comprehensive and systematic sufficient conditions for uniqueness of the Kantorovitch optimizer, and to exhibit continuous costs on arbitrary manifolds for which optimal plans are unique, despite the fact that such plans are not generally concentrated on graphs. We shall also establish a practical criterion for the uniqueness of the Kantorovitch optimizer in the non-compact setting on Polish spaces.

Contents

1 Introduction **2**

2 Sufficient conditions for the uniqueness **5**

2.1 Rooting sets and numbered limb systems 6

2.2 Wandering sets 8

2.3 c-extreme sets 9

3 Applications and Examples **12**

3.1 Sub-twist condition 12

3.2 Quadratic cost for pairs of small circles 13

4 Proofs of the results stated in Section 1 **18**

4.1 Proof of Theorem 1.4 18

4.2 Proof of Theorem 1.5 20

*This project was initiated at the Fields institute during the thematic program on Variational Problems in Physics, Economics and Geometry in 2014. The authors would like to thank the Fields institute for their hospitality. A. M. is grateful to Professor Robert McCann for many fruitful discussions on this work during the thematic program.

[†]School of Mathematics and Statistics, Carleton University, Ottawa, Ontario, K2C3L5, Canada (momeni@math.carleton.ca)

[‡]Université Nice Sophia Antipolis, Labo. J.-A. Dieudonné, UMR CNRS 7351, Parc Valrose, 06108 Nice Cedex 02, France (Ludovic.Rifford@math.cnrs.fr)

5 Proofs of the results stated in Section 2 **32**
5.1 Proof of Theorem 2.6 32
5.2 Proof of Theorem 2.10 35
5.3 Proof of Theorem 2.12 36

6 Uniqueness of optimal plans for infinite chains and non-compact settings **41**
6.1 Infinite chains and uniqueness 42
6.2 Quadratic cost on nested strictly convex sets 47

A Reminder on universally measurable sets **51**

B A description for (c, μ, ν)-minimizing sets **52**

C Partial numbered limb systems **58**

1 Introduction

Let M and N be smooth closed manifolds (meaning compact, without boundary) of dimensions m and $n \geq 1$ respectively, and $c : M \times N \to \mathbb{R}$ a continuous cost function. Given two Borel probability measures μ and ν on M and N, the Kantorovitch problem consists in minimizing the transportation cost

$$\int_{M \times N} c(x, y)\, d\gamma(x, y),$$

among all transport plans between μ and ν, meaning γ belongs to the set $\Pi(\mu, \nu)$ of Borel probability measures having marginals μ and ν. By compactness of M and N and continuity of c, minimizers of the Kantorovitch problem always exist [20]. In the present paper, we shall investigate sufficient conditions for uniqueness of solutions of the Kantorovitch problem, and illustrate continuous costs on arbitrary manifolds for which optimal plans are unique, even though such plans may not be generally concentrated on graphs.

Definition 1.1 (Uniquely minimizing costs). *A continuous cost $c : M \times N \to \mathbb{R}$ is called uniquely minimizing for the Kantorovitch problem, or simply uniquely minimizing, if for every pair of probability measures μ and ν respectively on M and N which are absolutely continuous with respect to the Lebesgue measures on M and N, the solution to the Kantorovitch problem is unique.*

Fix a continuous cost $c : M \times N \to \mathbb{R}$, given two Borel probability measures μ, ν respectively on M and N, it is well-known [18, 20] that there is a c-cyclically monotone compact set $\mathcal{S} \subset M \times N$ such that a transport plan $\gamma \in \Pi(\mu, \nu)$ is a solution to the Kantorovitch problem if and only if $\gamma(\mathcal{S}) = 1$. In the sequel, such a compact set $\mathcal{S}$ will be refered to a (c, μ, ν)-minimizing set. So, proving that c is uniquely minimizing for the Kantorovitch problem amounts to show that any (c, μ, ν)-minimizing set is a set of uniqueness with marginals μ, ν which means that it carries only one Borel probability measure on $M \times N$ with marginals μ, ν. Elaborating on previous works by Hestir and Williams [9], McCann and the second author [14] highlighted some properties on costs of class C^1 which are sufficient to insure that all (c, μ, ν)-minimizing sets are sets of uniqueness (provided the marginals are absolutely continuous with respect to Lebesgue measures). The aim of the present paper is twofold. Firstly, we extend the results in [14] to the case of continuous costs and to infer that uniquely minimizing costs are dense in the C^0 topology. Secondly, we shall establish some practical sufficient conditions for uniqueness applicable to a large class of optimal mass transport problems.

In order to define the notion of chains which is relevant in this paper, we need to introduce the concept of approximate sub-differential. Given a continuous function $f : M \to \mathbb{R}$ (or $f : N \to \mathbb{R}$) and $x \in M$, we call approximate sub-differential of f at x, denoted by $\tilde{D}^- f(x)$, the set of linear forms $p \in T_x^* M$ such that there are a function $\varphi : M \to \mathbb{R}$ of class C^1 and a Lebesgue measurable set $E \subset M$ which has Lebesgue density 1 at x such that

$$\varphi(x') \leq f(x') \quad \forall x' \in E, \quad \varphi(x) = f(x) \quad \text{and} \quad d_x\varphi = p.$$

In the case where f is differentiable at x we have $\tilde{D}^- f(x) = \{d_x f\}$. In the sequel, given a continuous cost $c : M \times N \to \mathbb{R}$, for every $(x, y) \in M \times N$, we denote by $\tilde{D}_x^- c(x, y)$ (resp. $\tilde{D}_y^- c(x, y)$) the approximate sub-differential of the function $x \mapsto c(x, y)$ at x (resp. $y \mapsto c(x, y)$ at y). The notion of chains introduced in [14] can be generalized as follows.

Definition 1.2 ($(c, \mathcal{S})$-chains and $(c, \mathcal{S})$-orbits). *Given a continuous cost $c : M \times N \to \mathbb{R}$ and a set $\mathcal{S} \subset M \times N$, we call $(c, \mathcal{S})$-chain of length $L \geq 2$, or simply $\mathcal{S}$-chain if $c \equiv 0$ and c-chain if $\mathcal{S} = M \times N$, any ordered family of pairs*

$$\Big((x_1, y_1), \ldots, (x_L, y_L)\Big) \in \mathcal{S}^L$$

such that the set

$$\Big\{(x_1, y_1), \ldots (x_L, y_L)\Big\}$$

is c-cyclically monotone and for every $l = 1, \ldots, L-1$ there holds, either

$$x_l = x_{l+1} \quad \textit{and} \quad y_l \neq y_{l+1} = y_{\min\{L, l+2\}} \quad \textit{and} \quad \tilde{D}_x^- c(x_l, y_l) \cap \tilde{D}_x^- c(x_l, y_{l+1}) \neq \emptyset,$$

or

$$y_l = y_{l+1} \quad \textit{and} \quad x_l \neq x_{l+1} = x_{\min\{L,l+2\}} \quad \textit{and} \quad \tilde{D}_y^- c(x_l, y_l) \cap \tilde{D}_y^- c(x_{l+1}, y_l) \neq \emptyset.$$

A $(c,\mathcal{S})$-chain of length $L \geq 5$ is called cyclic if there is $l \in \{5,\dots,L\}$ such that $(x_l,y_l) = (x_1,y_1)$, and the set $\mathcal{S}$ is said to be c-acyclic, or acyclic if $c \equiv 0$, if it contains no cyclic $(c,\mathcal{S})$-chain. Moreover, given $(x,y) \in \mathcal{S}$, we call $(c,\mathcal{S})$-orbit of (x,y), or simply $\mathcal{S}$-orbit if $c \equiv 0$ and c-orbit if $\mathcal{S} = M \times N$, the set of $(x',y') \in \mathcal{S}$ which can be joined to (x,y) through a $(c,\mathcal{S})$-chain, we denote it by $\mathcal{O}^{c,\mathcal{S}}(x,y)$ (or $\mathcal{O}^{\mathcal{S}}$ if $c \equiv 0$ and $\mathcal{O}^c$ if $\mathcal{S} = M \times N$). Lastly, given $(x,y) \in \mathcal{S}$ and an integer $L \geq 2$ we denote by $\mathcal{O}^{c,\mathcal{S}}_{\geq L}(x,y)$ (or $\mathcal{O}^{\mathcal{S}}_{\geq L}$ if $c \equiv 0$ and $\mathcal{O}^c_{\geq L}$ if $\mathcal{S} = M \times N$) the set of pairs $(x',y') \in \mathcal{S}$ for which there is a $(c,\mathcal{S})$-chain of length at least L from (x,y) to (x',y').

We observe that if both $\tilde{D}_x^- c(x,y)$ and $\tilde{D}_y^- c(x,y)$ are empty for some $(x,y) \in M \times N$ then (x,y) cannot belong to a $(c,\mathcal{S})$-subchain of length $L \geq 2$.

Definition 1.3 ($\mathcal{S}$-wandering costs). *Given a Borel set $\mathcal{S} \subset M \times N$, a continuous cost $c : M \times N \to \mathbb{R}$ is called $\mathcal{S}$-wandering, or simply wandering if $\mathcal{S} = M \times N$, if there are a partition $\{\mathcal{S}_i\}_{i \in \mathbb{N}}$ of $\mathcal{S}$ into countably many Borel sets and a sequence $\{L_i\}_{i \in \mathbb{N}}$ of integers ≥ 2 such that*

$$\mathcal{O}^{c,\mathcal{S}}_{\geq L_i}(x,y) \cap \mathcal{S}_i = \emptyset \qquad \forall (x,y) \in \mathcal{S}_i,\ \forall i \in \mathbb{N}.$$

We check easily that if a cost c is $\mathcal{S}$-wandering for a given set $\mathcal{S}$ then the set must be c-acyclic. The property of being wandering defined above is reminiscent to what happens in dynamics [11]. We shall show that the property of wanderingness, which is more general that the finiteness of chains as in [14], is sufficient for uniqueness. This is the content of Theorem 1.4. The possibility to work with Lipschitz costs will allow us to infer the density of uniquely minimizing costs in the C^0 topology in Theorem 1.5 below.

Given a pair of probability measures μ, ν respectively on M and N, we call (μ,ν)-null set in $M \times N$ any Borel set $\mathcal{N} \subset M \times N$ for which there are Borel sets $\mathcal{N}^1, \mathcal{N}^2 \subset M \times N$ with $\mathcal{N} = \mathcal{N}^1 \cup \mathcal{N}^2$ and $\pi^1(\mathcal{N}^1)$ has μ measure zero in M, and $\pi^2(\mathcal{N}^2)$ has ν measure zero in N (here and in the sequel $\pi^1 : M \times N \to M, \pi^2 : M \times N \to N$ stand for the projections onto the first and second variable). We call $(\mathcal{L}^m, \mathcal{L}^n)$-null set in $M \times N$ any set which is null as above with respect to the Lebesgue measures on M and N. The following theorem extends a previous result by McCann and the second author [14].

Theorem 1.4. *Let $c : M \times N \to \mathbb{R}$ be a continuous cost, μ, ν Borel probability measures respectively on M and N which are absolutely continuous with respect to the corresponding Lebesgue measures, and $\mathcal{S} \subset M \times N$ a (c,μ,ν)-minimizing set. Assume that there* *is a (μ,ν)-null set $\mathcal{N} \subset M \times N$ such that*

$$Lip_x^+ c(x,y) := \limsup_{\substack{x' \to x \\ x' \neq x}} \frac{c(x',y) - c(x,y)}{d^M(x,x')} < +\infty \tag{1.1}$$

and

$$Lip_y^+ c(x,y) := \limsup_{\substack{y' \to y \\ y' \neq y}} \frac{c(x,y') - c(x,y)}{d^N(y,y')} < +\infty \tag{1.2}$$

for every $(x,y) \in M \times N \setminus \mathcal{N}$, then there is a (μ,ν)-null-set $\mathcal{N}' \subset M \times N$ with $\mathcal{N} \subset \mathcal{N}'$ such that any $(0, \mathcal{S} \setminus \mathcal{N}')$-chain is a $(c, \mathcal{S} \setminus \mathcal{N}')$-chain. Moreover if c is $\mathcal{S} \setminus \mathcal{N}$-wandering, then there is a unique $\gamma \in \Pi(\mu,\nu)$ such that $\gamma(\mathcal{S}) = 1$.

Theorem 1.4 allows us to approximate any continuous cost by a continuous cost which enjoys the property of being uniquely minimizing.

Theorem 1.5 (C^0-density of uniquely minimizing costs)**.** *Let M and N be smooth closed manifolds of dimensions $n \geq 1$ and $c : M \times N \to \mathbb{R}$ a continuous cost function. Then for every $\epsilon > 0$, there is a uniquely minimizing cost $\tilde{c} : M \times N \to \mathbb{R}$ such that $\|\tilde{c} - c\|_{C^0} < \epsilon$.*

The proof of Theorem 1.5 consists in constructing a continuous cost with finite chains. The construction relies on a tiling of $M \times N$ associated with costs given by small perturbations of functions which are affine in charts. The idea of our proof does not allow to get the density for higher topologies (C^k or C^∞). We do not know if Theorem 1.5 holds in the C^1-topology.

The paper is organized as follows: Section 2 is concerned with sufficient conditions for the uniqueness of a plan supported on a given set. The new notions of wandering and c-extreme sets and corresponding sufficient results for uniqueness are introduced in Sections 2.2 and 2.3. They are illustrated by several examples in Section 3. The proofs of Theorems 1.4 and 1.5 are given in Section 4. Section 5 contains the proofs of the results stated in Section 2. We investigate the uniqueness problem in the non-compact setting in Section 6. Finally, several notions and results of importance for our proofs are recalled and discussed in the Appendix.

2 Sufficient conditions for the uniqueness

In the present section, we present a series of conditions on a given set to be the support of at most one probability measure with given marginals. We start by recalling classical results by Hestir and Williams [9] in the first paragraph, then we present a condition introduced by the first author [15] and we finish with the notion of wandering sets where we state the result which is required to prove the second part of Theorem 1.4. The proofs of all new results are given in Section 5.

2.1 Rooting sets and numbered limb systems

Let $\mathcal{S} \subset M \times N$ be fixed, we recall that for every $(x, y) \in \mathcal{S}$, we call $\mathcal{S}$-orbit of (x, y), the set of $(x', y') \in \mathcal{S}$ which can be joined to (x, y) through a $\mathcal{S}$-chain, we denote it by $\mathcal{O}^{\mathcal{S}}(x, y)$.

Definition 2.1 ($\mathcal{S}$-rooting set). *We call $\mathcal{S}$-rooting set, any set $\mathcal{R} \subset \mathcal{S}$ which satisfies the following properties:*

(i) $\mathcal{S} = \bigcup_{(x,y)\in\mathcal{R}} \mathcal{O}^{\mathcal{S}}(x, y)$.

(ii) For every $(x, y), (x', y') \in \mathcal{R}$, $(x, y) \neq (x', y') \Rightarrow (x', y') \notin \mathcal{O}^{\mathcal{S}}(x, y)$.

By the axiom of choice, any set $\mathcal{S}$ admits a $\mathcal{S}$-rooting set. If $\mathcal{S} \subset M \times N$ is acyclic, then from a given $\mathcal{S}$-rooting set $\mathcal{R}$ we can construct what is called a numbered limb system. Define the maps $F^1, F^2 : \mathcal{P}(M \times N) \to \mathcal{P}(M \times N)$ by

$$F^1(A) := \left(\pi^1\right)^{-1}\left[\pi^1(A)\right] \quad \text{and} \quad F^2(A) = \left(\pi^2\right)^{-1}\left[\pi^2(A)\right] \qquad \forall A \subset M \times N.$$

Then let for any $k \in \mathbb{N}^*$ (in the paper $\mathbb{N}$ stands for the set of nonnegative integers and $\mathbb{N}^*$ for the set of positive integers)

$$\begin{array}{rcl}
\mathcal{R}_1 & := & F^2(\mathcal{R}) \cap \mathcal{S} \\
\mathcal{R}_2 & := & F^1(\mathcal{R}_1) \cap \mathcal{S} \\
 & \vdots & \\
\mathcal{R}_{2k-1} & := & F^2(\mathcal{R}_{2k-2}) \cap \mathcal{S} \\
\mathcal{R}_{2k} & := & F^1(\mathcal{R}_{2k-1}) \cap \mathcal{S}.
\end{array}$$

For a map f from M to N, we denote by $\mathrm{Dom}(f)$ the domain of f, by $\mathrm{Ran}(f)$ its range and by $\mathrm{Graph}(f)$ its graph. If g is a map from N to M, then we denote by $\mathrm{Dom}(g)$ and $\mathrm{Ran}(g)$ its domain and range and we define its antigraph denoted by $\mathrm{Antigraph}(g)$ as the set of $(x, y) \in M \times N$ with $x = g(y)$ and $y \in \mathrm{Dom}(g)$. By acyclicity and the fact that $\mathcal{R}$ is an $\mathcal{S}$-rooting set, it can be shown that each $\mathcal{R}_k$ is the union of a graph and an antigraph (see the proof of Lemma 15 in [9]). For example, $\mathcal{R}_1$ is the union of $\mathcal{R}$ which is a graph (and an antigraph) and another set which is an antigraph thanks to (ii). Actually, the sets $\mathcal{R}_k$ form a so-called numbered limb system.

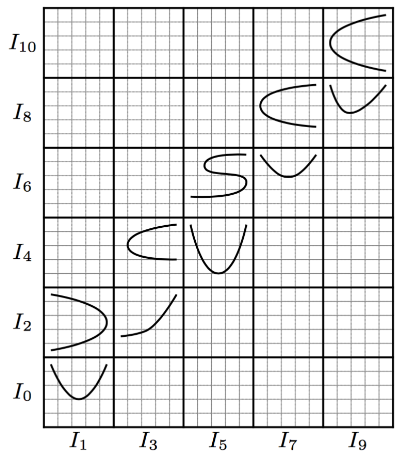

Figure 1: A numbered limb system

Definition 2.2 ($\mathcal{S}$-numbered limb system). *The set $\mathcal{S} \subset M \times N$ is said to admit a numbered limb system if there are countable disjoint decompositions of M and N,*

$$M = \bigcup_{i=0}^{\infty} I_{2i+1} \quad \textit{and} \quad N = \bigcup_{i=0}^{\infty} I_{2i},$$

with sequences of maps $\{f_{2i+1}\}_{i\in\mathbb{N}}$ and $\{f_{2i+2}\}_{i\in\mathbb{N}}$ of the form

$$f_{2i+1} : Dom(f_{2i+1}) \subset I_{2i+1} \subset M \longrightarrow Ran(f_{2i+1}) \subset I_{2i} \subset N$$

and

$$f_{2i+2} : Dom(f_{2i+2}) \subset I_{2i+2} \subset N \longrightarrow Ran(f_{2i+2}) \subset I_{2i+1} \subset M$$

such that

$$\mathcal{S} = \bigcup_{i=0}^{\infty} \Big(Graph(f_{2i+1}) \cup Antigraph(f_{2i+2}) \Big). \tag{2.1}$$

Each set $Graph(f_{2i+1})$ and $Antigraph(f_{2i+2})$ with $i \in \mathbb{N}$ is called a limb of $\mathcal{S}$.

The acyclicity and the existence of a rooting set or the existence of a numbered limb system is necessary for a set to carry only one probability measure with given marginals. The following is an easy consequence of Theorem 9 in [9].

Theorem 2.3. *Let $\mathcal{S} \subset M \times N$, the following properties are equivalent:*

(i) There exists an $\mathcal{S}$-rooting set and $\mathcal{S}$ is acyclic.

(ii) There exists an $\mathcal{S}$-numbered limb system.

Furthermore, if $\mathcal{S}$ is a Borel set in $M \times N$ and μ, ν are Borel probability measures respectively on M and N, and if there is exactly one measure $\gamma \in \Pi(\mu, \nu)$ such that $\gamma(\mathcal{S}) = 1$, then there is a Borel set $\mathcal{N} \subset M \times N$ with $\gamma(\mathcal{N}) = 0$ such that the set $\mathcal{S}' := \mathcal{S} \setminus \mathcal{N}$ is acyclic and admits an $\mathcal{S}'$-rooting set.

The converse holds only under additional measurability properties. The following is classical, see [9].

Theorem 2.4. *Let $\mathcal{S}$ be a Borel set in $M \times N$, μ, ν Borel probability measures respectively on M and N and $\gamma \in \Pi(\mu, \nu)$ with $\gamma(\mathcal{S}) = 1$, assume that one of the following properties is satisfied:*

(i) the set $\mathcal{S}$ is acyclic and there is a Borel $\mathcal{S}$-rooting set,

(ii) there exists a $\mathcal{S}$-numbered limb system with Borel limbs.

Then γ is the unique $\gamma' \in \Pi(\mu, \nu)$ such that $\gamma'(\mathcal{S}) = 1$.

Of course the above result holds if one of the assumptions (i) or (ii) is satisfied for a set $\mathcal{S}'$ of the form $\mathcal{S}' = \mathcal{S} \setminus \mathcal{N}$ with $\mathcal{N}$ a (μ, ν)-null set in $M \times N$. We turn now to other properties of chains which are sufficient for the uniqueness of probability measures (with given marginals) concentrated on a given set.

2.2 Wandering sets

The following notion is the counterpart of Definition 1.3 for sets.

Definition 2.5 (Wandering set). *A Borel set $\mathcal{S} \subset M \times N$ is called wandering if there are a countable partition of $\{\mathcal{S}_i\}_{i \in \mathbb{N}}$ of $\mathcal{S}$ into Borel sets and a sequence of integers $\{L_i\}_{i \in \mathbb{N}} \geq 2$ such that*

$$\mathcal{O}^{\mathcal{S}}_{\geq L_i}(x, y) \cap \mathcal{S}_i = \emptyset \qquad \forall (x, y) \in \mathcal{S}_i, \, \forall i \in \mathbb{N}.$$

In other words, a set is wandering if the null cost $c \equiv 0$ is $\mathcal{S}$-wandering. The existence of a "non-Borel" wandering set is necessary for uniqueness (it follows easily from Theorem 2.3), the converse holds.

Theorem 2.6. *Let $\mathcal{S}$ be a Borel set in $M \times N$, μ, ν Borel probability measures respectively on M and N and $\gamma \in \Pi(\mu, \nu)$ with $\gamma(\mathcal{S}) = 1$, assume that $\mathcal{S}$ is wandering. Then γ is the unique $\gamma' \in \Pi(\mu, \nu)$ such that $\gamma'(\mathcal{S}) = 1$.*

The proof of this theorem is given in Section 5.1. The second part of Theorem 1.4 is a straightforward application of Theorem 2.6. Let us now turn to another sufficient condition for uniqueness which does not rely on the notion of chains.

2.3 c-extreme sets

The results of uniqueness for c-extreme sets will follow from a result on Borel disjoint union of graphs and antigraph obtained by the first author in [15].

Definition 2.7 (Borel strongly disjoint union of a graph and an antigraph). *The set $\mathcal{S} \subset M \times N$ is said to be the strongly disjoint union of a graph and an antigraph if there are two functions*

$$f : Dom(f) \subset M \longrightarrow N \quad \text{and} \quad g : Dom(g) \subset N \longrightarrow M$$

which satisfy the following properties:

(i) $Graph(f) \cap Antigraph(g) = \emptyset$,

(ii) there exists a bounded function $\theta : N \to \mathbb{R}$ such that $\theta(f \circ g(y)) > \theta(y)$ for every $y \in Dom(f \circ g)$.

Furthermore, if the functions f, g and θ are Borel measurable then the set $\mathcal{S}$ is said to be the Borel strongly disjoint union of a graph and an antigraph.

Borel strongly disjoint union of a graph and an antigraph is reminiscent to the notion of aperiodic measurable decomposition introduced by Beneš and Štěpán [3]. However it is more practical and easier to handle. The following theorem is proven in [15].

Theorem 2.8. *Let $\mathcal{S}$ be a Borel set in $M \times N$, μ, ν Borel probability measures respectively on M and N and $\gamma \in \Pi(\mu, \nu)$ with $\gamma(\mathcal{S}) = 1$. The following assertions hold:*

(i) Assume that $\mathcal{S}$ is the Borel strongly disjoint union of a graph and an antigraph. Then γ is the unique $\gamma' \in \Pi(\mu, \nu)$ such that $\gamma'(\mathcal{S}) = 1$.

(ii) Assume that γ is the unique $\gamma' \in \Pi(\mu, \nu)$ such that $\gamma'(\mathcal{S}) = 1$. Then the set is $\mathcal{S}$ strongly disjoint union of a graph and an antigraph.

Let us now introduce the notion of c-extreme sets. Let $c : M \times N \to \mathbb{R}$ be a continuous cost, μ, ν Borel probability measures respectively on M and N, and $\mathcal{S} \subset M \times N$ a (c, μ, ν)-minimizing set. We define the set-valued function $\Gamma_{\mathcal{S}} : M \to 2^N$ by

$$\Gamma_{\mathcal{S}}(x) = \Big\{ y \in N \mid (x, y) \in \mathcal{S} \Big\},$$

with $Dom(\Gamma_{\mathcal{S}}) = \{x \in M;\, \Gamma_{\mathcal{S}}(x) \neq \emptyset\}$. Note that by compactness of $\mathcal{S}$ for every $x \in M$ the set $\Gamma_{\mathcal{S}}(x)$ is a compact subset of N and the graph of $\Gamma_{\mathcal{S}}$ is a compact subset of $M \times N$. Then we define the map $f_{\mathcal{S},c} : M \to 2^N$ by

$$f_{\mathcal{S},c}(x) := \operatorname{argmax}\Big\{c(x,y) \,|\, y \in \Gamma_{\mathcal{S}}(x)\Big\} = \Big\{y \in \Gamma_{\mathcal{S}}(x) \,|\, c(x,y) = \max_{z\in\Gamma_{\mathcal{S}}(x)} c(x,z)\Big\}.$$

We can now introduce the notion of c-extreme sets.

Definition 2.9 (c-extreme sets). *We say that a set $\mathcal{S} \subset M \times N$ is c-extreme if there are full measure Borel sets $\bar{M}$ of M and $\bar{N}$ of N such that for all distinct points $x_1, x_2 \in \bar{M}$ the following assertion holds:*

$$\Big\{\Gamma_{\mathcal{S}}(x_1) \setminus \{y_1\}\Big\} \cap \Big\{\Gamma_{\mathcal{S}}(x_2) \setminus \{y_2\}\Big\} \cap \bar{N} \quad = \quad \emptyset,$$

for all $y_1 \in f_{\mathcal{S},c}(x_1)$ and $y_2 \in f_{\mathcal{S},c}(x_2)$.

Basically, the set $\mathcal{S}$ is c-extreme if the set $\big(\mathcal{S} \setminus \mathrm{Graph}(f_{\mathcal{S},c|D(\mathcal{S},c)})\big) \cap \big(\bar{M} \times \bar{N}\big)$ is nothing but the graph of a function from N to M. Here,

$$D(\mathcal{S},c) := \Big\{x \in M \,|\, \{\text{argmax } c(x,y);\, y \in \Gamma_{\mathcal{S}}(x)\} \text{ is a singleton } \Big\},$$

and $f_{\mathcal{S},c|D(\mathcal{S},c)}$ is the restriction of the map to the set $D(\mathcal{S},c)$. The proof of the following result is given in Section 5.2.

Theorem 2.10. *Let $c : M \times N \to \mathbb{R}$ be a continuous cost, μ, ν Borel probability measures respectively on M and N, and $\mathcal{S} \subset M \times N$ a (c,μ,ν)-minimizing set. If $\mathcal{S}$ is c-extreme then there is a unique $\gamma \in \Pi(\mu,\nu)$ such that $\gamma(\mathcal{S}) = 1$.*

Definition 2.9 of $c-$extreme sets is too strong in some cases. The purpose of the next theorem is to show that a local version of Definition 2.9 suffices to prove uniqueness.

A family $P = \{Y_i\}_{i\in I}$ with $I \subseteq \mathbb{N}$ is called an ordered partition of N provided the sets Y_i with $i \in I$ are pairwise disjoint and their union is N. It is called a Borel ordered partition if all the Y_i's (with $i \in I$) are Borel subsets of N. Let $\mathcal{S} \subset M \times N$ be a (c,μ,ν)-minimizing set. Given an ordered partition $P = \{Y_i\}_i$ of N, we set

$$D(\mathcal{S},c,P) := \Big\{x \in M \,|\, \{\text{argmax } c(x,y);\, y \in \Gamma_{\mathcal{S}}(x) \cap Y_{i(x)}\} \text{ is a singleton } \Big\},$$

where the function $i : \mathrm{Dom}(\Gamma_{\mathcal{S}}) \to \mathbb{N}$ is given by

$$i(x) := \min\Big\{i \in \mathbb{N} \,|\, \Gamma_{\mathcal{S}}(x) \cap Y_i \neq \emptyset\Big\},$$

and we define the map $f_{\mathcal{S},c,P} : M \to 2^N$ by

$$f_{\mathcal{S},c,P}(x) := \operatorname{argmax}\Big\{c(x,y) \,|\, y \in \Gamma_{\mathcal{S}}(x) \cap Y_{i(x)}\Big\}.$$

The following definition can be seen as a local version of Definition 2.9.

Definition 2.11 ((c,P)-extreme sets). *Let $P = \{Y_i\}_{i\in I\subseteq\mathbb{N}}$ be a Borel ordered partition of N. We say that $\mathcal{S}$ is (c,P)-extreme if there are full measure Borel sets $\bar{M}$ of M and $\bar{N}$ of N such that the following assertions hold:*

(i) for each $x \in Dom\big(\Gamma_{\mathcal{S}}\big) \cap \bar{M}$ the following set is non-empty,

$$\Big\{ \operatorname{argmax} c(x,y);\ y \in \Gamma_{\mathcal{S}}(x) \cap Y_{i(x)} \Big\};$$

(ii) for all distinct points $x_1, x_2 \in \bar{M}$,

$$\Big\{\Gamma_{\mathcal{S}}(x_1) \setminus \{y_1\}\Big\} \cap \Big\{\Gamma_{\mathcal{S}}(x_2) \setminus \{y_2\}\Big\} \cap \bar{N} \;\;=\;\; \emptyset,$$

for all $y_1 \in f_{\mathcal{S},c,P}(x_1)$ and $y_2 \in f_{\mathcal{S},c,P}(x_2)$.

If P is the trivial partition $P = \{N\}$, then (c,P)-extreme sets coincide with c-extreme sets defined in Definition 2.9. We have the following result (whose proof is given in Section 5.3).

Theorem 2.12. *Let $c : M \times N \to \mathbb{R}$ be a continuous cost, μ, ν Borel probability measures respectively on M and N, and $\mathcal{S} \subset M \times N$ a (c,μ,ν)-minimizing set. If there exists a Borel ordered partition P of N such that $\mathcal{S}$ is (c,P)-extreme, there is a unique $\gamma \in \Pi(\mu,\nu)$ such that $\gamma(\mathcal{S}) = 1$.*

Before going to applications, let us recall the following standard description result about (c,μ,ν)-minimizing sets (see for example [18, 20]).

Lemma 2.13. *Let M and N be smooth closed manifolds. Let $c : M \times N \to \mathbb{R}$ be a continuous cost, μ, ν Borel probability measures respectively on M and N, $\mathcal{S} \subset M \times N$ a (c,μ,ν)-minimizing set. Then there are continuous potentials $\psi : M \to \mathbb{R}$ and $\phi : N \to \mathbb{R}$ which satisfy*

$$\psi(x) = \max\Big\{\phi(y) - c(x,y) \,|\, y \in N\Big\} \qquad \forall x \in M,$$

$$\phi(y) = \min\Big\{\psi(x) + c(x,y) \,|\, x \in M\Big\} \qquad \forall y \in N,$$

$$\textit{and} \quad \mathcal{S} \subset \partial_c\psi := \Big\{(x,y) \in M \times N \,|\, c(x,y) = \phi(y) - \psi(x)\Big\}.$$

Note also that by compactness and continuity of c, the set $\partial_c\psi$ is compact and all its fibers over M and N are nonempty and compact.

3 Applications and Examples

We shall now proceed with some applications of our results in this section.

3.1 Sub-twist condition

Here we provide a new proof for the following theorem originally established in [1].

Theorem 3.1. *Let M and N be smooth closed manifolds equipped with Borel probability measures μ on M and ν on N . Let $c \in C^1(M \times N)$ satisfy the sub-twist condition, that is, for each $y_1 \neq y_2 \in N$ the map*

$$x \in M \to c(x, y_1) - c(x, y_2)$$

has no critical points, save at most one global minimum and at most one global maximum. If μ is absolutely continuous in each coordinate chart on N then any (c, μ, ν)-minimizing set is a set of uniqueness.

Proof. Let $\mathcal{S}$ be a (c, μ, ν)-minimizing set. We show that $\mathcal{S}$ is c-extreme as in Definition 2.9. Without loss of generality we can assume that there exists $\bar{x} \in M$ such that $c(\bar{x}, y) = 0$ for all $y \in N$. In fact, one can always replace c by the function $\bar{c}$ defined by $\bar{c}(x, y) = c(x, y) - c(\bar{x}, y)$ where $\bar{x} \in M$ is fixed. By Lemma 2.13, there exist continuous functions $\psi : M \to \mathbb{R}$ and $\phi : N \to \mathbb{R}$ with

$$\phi(y) = \min_{x \in M}\{c(x, y) + \psi(x)\} \qquad \& \qquad \psi(x) = \max_{y \in N}\{\phi(y) - c(x, y)\}, \tag{3.1}$$

such that

$$\mathcal{S} \subseteq \big\{(x, y) \in M \times N;\ \phi(y) - \psi(x) = c(x, y)\big\}.$$

Note that (3.1) implies that ψ is locally Lipschitz and therefore almost surely differentiable due to absolute continuity of μ. Set $\bar{M} = Dom(D\psi) \cap Spt(\mu) \setminus \{\bar{x}\}$. Note that for every $x \in \bar{M}$, $\{\operatorname{argmax} c(x, y);\ y \in \Gamma_{\mathcal{S}}(x)\}$ is a singleton. Indeed, if for some $x_0 \in \bar{M}$, there exist $y_1, y_2 \in \{\operatorname{argmax} c(x, y);\ y \in \Gamma_{\mathcal{S}}(x)\}$ then x_0 is a critical point of $x \to c(x, y_1) - c(x, y_2)$. If x_0 is a global minimum or maximum then $c(x, y_1) - c(x, y_2) \neq 0$ for all $x_0 \neq x \in X$. This leads to a contradiction as $c(\bar{x}, y_1) - c(\bar{x}, y_2) = 0$. This shows that $\bar{M} \subseteq D(\mathcal{S}, c)$. We show that

$$\big\{\Gamma_{\mathcal{S}}(x_1) \setminus \{f_{\mathcal{S},c}(x_1)\}\big\} \cap \Gamma_{\mathcal{S}}(x_2) = \emptyset, \tag{3.2}$$

for all distinct points $x_1, x_2 \in \bar{M}$. Take $y \in \Gamma_{\mathcal{S}}(x_2) \cap \big\{\Gamma_{\mathcal{S}}(x_1) \setminus \{f_{\mathcal{S},c}(x_1)\}\big\}$. It follows that x_1 is a critical point of the function

$$x \to c(x, y) - c\big(x, f_{\mathcal{S},c}(x_1)\big),$$

and since $c\big(x_1, f_{\mathcal{S},c}(x_1)\big) > c(x_1, y)$ we have that x_1 has to be a global minimum of the function

$$x \to c(x,y) - c\big(x, f_{\mathcal{S},c}(x_1)\big).$$

It then follows that

$$c(x_2,y) - c\big(x_2, f_{\mathcal{S},c}(x_1)\big) > c(x_1,y) - c\big(x_1, f_{\mathcal{S},c}(x_1)\big).$$

Therefore,

$$c(x_2,y) + c\big(x_1, f_{\mathcal{S},c}(x_1)\big) > c(x_1,y) + c\big(x_2, f_{\mathcal{S},c}(x_1)\big),$$

which contradicts the fact that $\mathcal{S}$ is $c-$cyclically monotone. Thus, $\mathcal{S}$ is c-extreme and the result follows from Theorem 2.10. □

Remark 3.2. *In most applications, the full measure set $\bar{M}$ in Definition 2.9 can be chosen so that $\bar{M} \subseteq D(\mathcal{S}, c)$. In this case a set $\mathcal{S}$ is c-extreme if and only if*

$$\Big\{\Gamma_{\mathcal{S}}(x_1) \setminus \{f_{\mathcal{S},c}(x_1)\}\Big\} \cap \Big\{\Gamma_{\mathcal{S}}(x_2) \setminus \{f_{\mathcal{S},c}(x_2)\}\Big\} = \emptyset, \qquad \forall x_1 \neq x_2 \in \bar{M}.$$

Let us compare the latter expression with previously known results for the uniqueness of optimal transportation on smooth manifolds:

- *Twist condition: In this case for all $x \in \bar{M}$, the set $f_{\mathcal{S},c}(x)$ is a singleton and therefore*

$$\Gamma_{\mathcal{S}}(x) \setminus \{f_{\mathcal{S},c}(x)\} = \emptyset, \qquad \forall x \in \bar{M}.$$

- *Sub-Twist condition: In this case, as shown in (3.2) within the proof of Theorem 3.1,*

$$\big\{\Gamma_{\mathcal{S}}(x_1) \setminus \{f_{\mathcal{S},c}(x_1)\}\big\} \cap \Gamma_{\mathcal{S}}(x_2) = \emptyset, \qquad \forall x_1 \neq x_2 \in \bar{M}.$$

Evidently, the notion of c-extreme sets seems to be a straightforward generalization of the twist-type conditions stated in Remark 3.2.

3.2 Quadratic cost for pairs of small circles

Let us consider two circles $\mathcal{S}_1, \mathcal{S}_2$ (see Figure 2) in the plane centered respectively at $(-1,0)$ and $(1,0)$ of radius $\rho \in (0,1)$. Let $M = \mathcal{S}_1 \cup \mathcal{S}_2$ and suppose that μ and ν are probability measures on M. Euclidean norm and the inner product in $\mathbb{R}^2$ are denoted by $|.|$ and $\langle ., . \rangle$ respectively. We shall consider the following problem,

$$\inf\Big\{\int_{M\times M} c(x,y)\, d\gamma;\ \gamma \in \Pi(\mu,\nu)\Big\}, \tag{3.3}$$

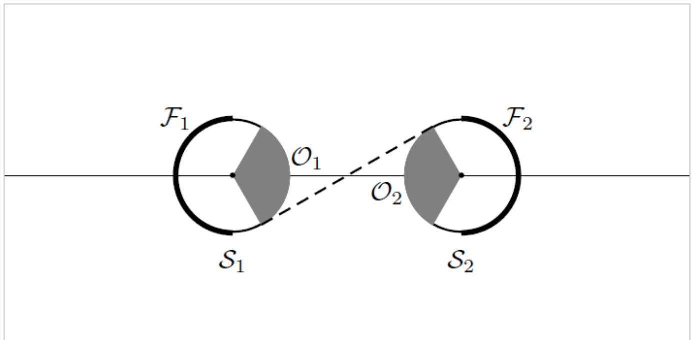

Figure 2: A pair of circles

where $c(x,y) = |x-y|^2$.

We need some preliminaries before stating our result. Define the exterior faces $\mathcal{F}_1, \mathcal{F}_2$ of M by

$$\mathcal{F}_1 := \Big\{ x \in \mathcal{S}_1 \,|\, \langle x - (-1,0), (1,0) \rangle < 0 \Big\}$$

$$\text{and} \quad \mathcal{F}_2 := \Big\{ x \in \mathcal{S}_2 \,|\, \langle x - (1,0), (1,0) \rangle > 0 \Big\}.$$

Define the interior faces $\mathcal{I}_1, \mathcal{I}_2$ by

$$\mathcal{I}_1 := \mathcal{S}_1 \setminus \mathcal{F}_1 \quad \text{and} \quad \mathcal{I}_2 := \mathcal{S}_2 \setminus \mathcal{F}_2.$$

At each point $z \in M$, let $n(z)$ be the unit outward normal to M at point z. Finally, define the two symmetric eyes $\mathcal{O}_1, \mathcal{O}_2$ by

$$\mathcal{O}_1 := \Big\{ z \in \mathcal{I}_1 \,|\, \exists z_1 \in \mathcal{S}_1, \exists s \in \mathbb{R} \text{ s.t. } z_1 + sn(z) \in \mathcal{S}_2 \Big\}$$

$$\text{and} \quad \mathcal{O}_2 := \Big\{ z \in \mathcal{I}_2 \,|\, \exists z_2 \in \mathcal{S}_2, \exists s \in \mathbb{R} \text{ s.t. } z_2 + sn(z) \in \mathcal{S}_1 \Big\}.$$

We have the following result for the uniqueness.

Theorem 3.3. *Let $M = \mathcal{S}_1 \cup \mathcal{S}_2$ and, let $c : M \times M \to \mathbb{R}$ be defined by $c(x,y) = |x-y|^2$. Assume that probability measures μ and ν on M are absolutely continuous in each coordinate chart on M. If $\rho < 1/2$ and $\nu(\mathcal{I}_1 \setminus \mathcal{O}_1) = 0$, then any (c,μ,ν)-minimizing set $\mathcal{S}$ is a set of uniqueness.*

We shall need some preliminary lemmas before proving Theorem 3.3. Let $\mathcal{S}$ be a (c,μ,ν)-minimizing set.

Lemma 3.4. *There exist a full ν-measure subset N_0 of M and a full μ-measure subset M_0 of M such that:*
1. for each $y \in N_0$ if there exist $x_1, x_2 \in M$ with (x_1, y) and $(x_2, y) \in \mathcal{S}$ then there exists $\alpha \in \mathbb{R}$ such that $n(y) = \alpha(x_1 - x_2)$.
2. for each $x \in M_0$ if there exist $y_1, y_2 \in M$ with (x, y_1) and $(x, y_2) \in \mathcal{S}$ then there exists $\beta \in \mathbb{R}$ such that $n(x) = \alpha(y_1 - y_2)$.

Proof. Let $M_0 = Dom(D\psi)$ and $N_0 = Dom(D\phi)$ where ϕ and ψ are given in Lemma 2.13. Now for each $y \in N_0$ if there exist $x_1, x_2 \in M$ with (x_1, y) and $(x_2, y) \in \mathcal{S}$ then we must have $D_y(x_1, y) = D_y(x_2, y)$ from which the result follows. Proof of the second part is similar. □

Let $\mathcal{T}$ be the set of north and south poles of M;

$$\mathcal{T} = \big\{(-1, -\rho), (-1, \rho), (1, -\rho), (1, \rho)\big\}.$$

Since $\nu(\mathcal{I}_1 \setminus \mathcal{O}_1) = \nu(\mathcal{T}) = \mu(\mathcal{T}) = 0$, it follows from Lemma 3.4 that the sets

$$\bar{M} = M_0 \setminus \mathcal{T} \quad \& \quad \bar{N} = N_0 \setminus \Big(\mathcal{T} \cup (\mathcal{I}_1 \setminus \mathcal{O}_1)\Big) \tag{3.4}$$

are μ and ν full measure subsets of M respectively. By setting $Y_1 = \mathcal{S}_2$, $Y_2 = \mathcal{F}_1 \cup \mathcal{O}_1$ and $Y_3 = \mathcal{S}_1 \setminus Y_2$ we have that $P := \{Y_1, Y_2, Y_3\}$ is a Borel ordered partition of $\mathcal{S}$. We will use Theorem 2.12 to prove uniqueness.
In the following three lemmas, let $x_1, x_2 \in \bar{M}$, $y \in \bar{N}$ and $y_1, y_2 \in M$ be such that $(x_i, y_i), (x_i, y) \in \mathcal{S}$ for $i = 1, 2$.

Lemma 3.5. *The following assertions hold:*
1. $y \notin \mathcal{F}_1 \cup \mathcal{F}_2$.
2. y, y_1, y_2 can not belong to the same circle.

Proof. If $y \in \mathcal{F}_1 \cup \mathcal{F}_2$ or y, y_1, y_2 belong to the same circle then

$$\langle n(y), y - y_1\rangle > 0 \qquad \& \qquad \langle n(y), y - y_2\rangle > 0.$$

It follows from Lemma 3.4 that $n(y) = \alpha(x_2 - x_1)$ for some $\alpha \in \mathbb{R}$. Plugging this into above inequalities we get

$$\alpha\langle x_2 - x_1, y - y_1\rangle > 0 \qquad \& \qquad \alpha\langle x_2 - x_1, y - y_2\rangle > 0.$$

This leads to a contradiction since by the $c-$monotonicity of the set $\mathcal{S}$ we must have

$$\langle x_2 - x_1, y - y_1\rangle > 0 \qquad \& \qquad \langle x_2 - x_1, y - y_2\rangle < 0.$$

□

Lemma 3.6. *If either of the following conditions holds then* $y \in \mathcal{F}_1 \cup \mathcal{F}_2$*,*
1. $x_2, y_1 \in \mathcal{S}_1$ *and* $x_1, y_2 \in \mathcal{S}_2$*;*
2. $x_2, y_1 \in \mathcal{S}_2$ *and* $x_1, y_2 \in \mathcal{S}_1$*.*

Proof. We prove this Lemma under Condition 1). The other case is similar. Without loss of generality we can assume that $y \in \mathcal{S}_2$. It follows from Lemma 3.4 that $n(y) = \alpha(x_2 - x_1)$ for some $\alpha \in \mathbb{R}$. Since $y, y_2 \in \mathcal{S}_2$, we obtain that $\langle n(y), y - y_2 \rangle > 0$. It then follows that

$$\alpha \langle x_2 - x_1, y - y_2 \rangle > 0.$$

This together with the $c-$monotonicity of the set $\mathcal{S}$ imply that $\alpha < 0$. Assuming $x_1 - x_2 = r\epsilon^{i\theta}$, it yields that $n(y) = \epsilon^{i\theta}$. Since $x_2 \in S_1$ and $x_1 \in S_2$, it follows that

$$-\arctan(\frac{\rho}{1-\rho}) \leq \theta \leq \arctan(\frac{\rho}{1-\rho}).$$

Since $\rho < 1/2$ we have that $\arctan(\frac{\rho}{1-\rho}) < \pi/4$. Therefore $-\frac{\pi}{4} < \theta < \frac{\pi}{4}$, from which together with the fact that $y \in \mathcal{S}_2$ we must have $y \in \mathcal{F}_2$. $\square$

Lemma 3.7. *Let* $y_1 \in \mathcal{S}_i$ *and* $y_2 \in M \setminus \mathcal{S}_i$ *for* $i = 1$ *or* $i = 2$*. Furthermore, assume that*

$$|y_2 - x_2| \geq |y - x_2| \qquad \& \qquad |y_1 - x_1| \geq |y - x_1|.$$

The following assertions hold:
1. If $y \in M \setminus \mathcal{S}_i$ *then* $x_2 \in \mathcal{S}_i$ *and* $x_1 \in M \setminus \mathcal{S}_i$*.*
2. If $y \in \mathcal{S}_i$ *then* $x_1 \in M \setminus \mathcal{S}_i$ *and* $x_2 \in \mathcal{S}_i$*.*

Proof. We just prove part 1). A similar argument works for the second part. Since $y_1 \in \mathcal{S}_i$ and $y \in M \setminus \mathcal{S}_i$ we must have $x_1 \in M \setminus \mathcal{S}_i$. Indeed, if $x_1 \in \mathcal{S}_i$ then

$$|y_1 - x_1| \leq 2\rho < 2 - 2\rho = dist(S_1, S_2) \leq |y - x_1|, \qquad (\text{since } \rho < \frac{1}{2}),$$

which violates the hypothesis that $|y_1 - x_1| \geq |y - x_1|$. Now assume that $x_2 \in M \setminus \mathcal{S}_i$. Then it follows from $x_1, x_2 \in M \setminus \mathcal{S}_i$ that $\langle n(x_2), x_2 - x_1 \rangle > 0$. By Lemma 3.4 we have that $n(x_2) = \alpha(y - y_2)$ for some $\alpha \in \mathbb{R}$. Therefore, $\alpha \langle y - y_2, x_2 - x_1 \rangle > 0$ from which together with the $c-$monotonicity of the $\mathcal{S}$ we obtain that $\alpha < 0$. We also have that

$$\langle n(x_2), y_2 - x_2 \rangle \leq 0 \qquad \& \qquad \langle n(x_2), y - x_2 \rangle \leq 0,$$

where at least one of the above inequalities is strict. Thus,

$$\langle n(x_2), y_2 + y - 2x_2 \rangle < 0.$$

Substituting $n(x_2) = \alpha(y - y_2)$ in the latter inequality together with the fact that $\alpha < 0$ imply that

$$\langle y_2 - y, y_2 + y - 2x_2 \rangle < 0,$$

from which we obtain

$$|y_2 - x_2|^2 < |y - x_2|^2.$$

This is a contradiction as by the hypothesis we have that $|y_2 - x_2| \geq |y - x_2|$. $\square$

Completion of the proof of Theorem 3.3. We just need to show that $\mathcal{S}$ is $(c, P)-$extreme where $P = \{Y_1, Y_2, Y_3\}$. As in Definition 2.11, define the maps $i : Dom(\Gamma_{\mathcal{S}}) \to \{1, 2, 3\}$ and $f_{\mathcal{S},c,P} : M \to 2^M$ by

$$i(x) = \min\{i \in \mathbb{N}; \Gamma_{\mathcal{S}}(x) \cap Y_i \neq \emptyset\}$$

and

$$f_{\mathcal{S},c,P}(x) = \operatorname{argmax} \big\{ c(x, y);\ y \in \Gamma_{\mathcal{S}}(x) \cap Y_{i(x)} \big\}.$$

Note that for each $x \in \bar{M}$ the set $\Gamma_{\mathcal{S}}(x)$ is finite and therefore condition $i)$ in Definition 2.11 is satisfied. Take distinct points $x_1, x_2 \in \bar{M}$. Take $y_1 \in f_{\mathcal{S},c,P}(x_1)$ and $y_2 \in f_{\mathcal{S},c,P}(x_2)$. We need to show that

$$\big\{\Gamma_{\mathcal{S}}(x_1) \setminus \{y_1\}\big\} \cap \big\{\Gamma_{\mathcal{S}}(x_2) \setminus \{y_2\}\big\} \cap \bar{N} = \emptyset,$$

where $\bar{N}$ is the set given in (3.4). Take

$$y \in \big\{\Gamma_{\mathcal{S}}(x_1) \setminus \{y_1\}\big\} \cap \big\{\Gamma_{\mathcal{S}}(x_2) \setminus \{y_2\}\big\} \cap \bar{N}.$$

Since $Y_3 \cap \bar{N} = \emptyset$, we obtain that $y \notin Y_3$. If $y \in Y_1$, it follows from the definition of $f_{\mathcal{S},c,P}$ that $y_1, y_2 \in Y_1$ and this never happens due to Lemma 3.5. Thus, we assume that $y \in Y_2$. Note that both y_1, y_2 can not belong to Y_2 by virtue of Lemma 3.5. Without loss of generality we assume that $y_1 \in Y_2 \subset \mathcal{S}_1$ and $y_2 \in Y_1 = \mathcal{S}_2$. Note also that $y_1 = \operatorname{argmax}_{z \in \Gamma_{\mathcal{S}}(x_1)} |z - x_1|^2$ and $y_2 = \operatorname{argmax}_{z \in \Gamma_{\mathcal{S}}(x_2)} |z - x_2|^2$ from which we obtain that

$$|y_1 - x_1| \geq |y - x_1| \qquad \& \qquad |y_2 - x_2| \geq |y - x_2|.$$

Since $y \in \mathcal{S}_1$, it follows from Lemma 3.7 and the latter inequalities that $x_1 \in \mathcal{S}_2$ and $x_2 \in \mathcal{S}_1$. Therefore, we have that $y_1, x_2 \in \mathcal{S}_1$ and $y_2, x_1 \in \mathcal{S}_2$. Thus, by Lemma 3.6 we must have that $y \in \mathcal{F}_1$ which is impossible due to Lemma 3.5. This completes the proof. $\square$

4 Proofs of the results stated in Section 1

4.1 Proof of Theorem 1.4

For a (c,μ,ν)-minimizing set $\mathcal{S} \subset M \times N$, let continuous potentials ψ and ϕ be as in Lemma 2.13, i.e.,

$$\psi(x) = \max\Big\{\phi(y) - c(x,y) \,|\, y \in N\Big\} \qquad \forall x \in M, \tag{4.1}$$

$$\phi(y) = \min\Big\{\psi(x) + c(x,y) \,|\, x \in M\Big\} \qquad \forall y \in N, \tag{4.2}$$

$$\text{and} \quad \mathcal{S} \subset \partial_c\psi := \Big\{(x,y) \in M \times N \,|\, c(x,y) = \phi(y) - \psi(x)\Big\}. \tag{4.3}$$

Set

$$\partial_c\psi(x) := \Big\{y \in N \,|\, (x,y) \in \partial_c\psi\Big\} \qquad \forall x \in M,$$
$$\partial^c\phi(y) := \Big\{x \in M \,|\, (x,y) \in \partial_c\psi\Big\} \qquad \forall y \in N.$$

Let $\mathcal{N} = \mathcal{N}^1 \cup \mathcal{N}^2$ be a null set (with $\pi^1(\mathcal{N}^1)$ of Lebesgue measure zero in M and $\pi^2(\mathcal{N}^2)$ of Lebesgue measure zero in N) such that (1.1)-(1.2) hold on $M \times N \setminus \mathcal{N}$. Define the sets $\mathcal{E}^M \subset M, \mathcal{E}^N \subset N$ by

$$\mathcal{E}^M := \Big\{x \in M \,|\, (x,y) \in \mathcal{N},\, \forall y \in \partial_c\psi(x)\Big\}$$

and

$$\mathcal{E}^N := \Big\{y \in N \,|\, (x,y) \in \mathcal{N},\, \forall x \in \partial^c\phi(y)\Big\}.$$

By construction, the following holds (the notion of universally measurable set is recalled in Appendix A).

Lemma 4.1. *The sets $\mathcal{E}^M \subset M$ and $\mathcal{E}^N \subset N$ are universally measurable and satisfy $\mu(\mathcal{E}^M) = \nu(\mathcal{E}^N) = 0$.*

Proof of Lemma 4.1. Note that

$$\mathcal{E}^M = \pi^2(\mathcal{N}) \setminus \pi^2\left(\partial_c\psi \cap \mathcal{N}^c\right).$$

Let γ be an optimal transport plan between μ and ν. Then we have

$$\mathcal{S} \subset \big(\mathcal{S} \setminus (\mathcal{E}^M \times N)\big) \cup \mathcal{N} \quad \text{and} \quad \gamma(\mathcal{S}) = 1.$$

Moreover, since both μ and ν are absolutely continuous with respect to Lebesgue and $\mathcal{N}$ is a null set, we have $\gamma(\mathcal{N}) \leq \gamma(\mathcal{N}^1) + \gamma(\mathcal{N}^2) = 0$. Therefore, there holds

$$\mu(\mathcal{E}^M) = \gamma\left(\mathcal{E}^M \times N\right) = 0.$$

The proof of $\nu(\mathcal{E}^N) = 0$ follows the same lines. $\square$

Assumptions (1.1)-(1.2) allow to show that the potentials ψ and ϕ are approximately differentiable almost everywhere. We recall that a measurable function $f : M \to \mathbb{R}$ (or from N to $\mathbb{R}$) is approximately differentiable at $x \in M$ provided there is a function $\varphi : M \to \mathbb{R}$ of class C^1 and a Lebesgue measurable set $E \subset M$ which has Lebesgue density 1 at x such that $f(x') = \varphi(x')$ for all $x' \in E$.

Lemma 4.2. *There is a Borel set $\mathcal{F}^M \subset M$ (resp. $\mathcal{F}^N \subset N$) with $\mathcal{E}^M \subset \mathcal{F}^M$ and $\mu(\mathcal{F}^M) = 0$ (resp. with $\mathcal{E}^N \subset \mathcal{F}^N$ and $\nu(\mathcal{F}^N) = 0$) such that ψ (resp. ϕ) is approximately differentiable at every $x \in M \setminus \mathcal{F}^M$ (resp. $y \in M \setminus \mathcal{F}^N$).*

Proof of Lemma 4.2. By construction, for every $x \in M \setminus \mathcal{E}^M$ there is $y \in \partial_c\psi(x)$ such that $\mathrm{Lip}^+_x c(x,y) < +\infty$, which implies readily (by (4.1))

$$\mathrm{Lip}^-\psi(x) := \liminf_{\substack{x' \to x \\ x' \neq x}} \frac{\psi(x') - \psi(x)}{d^M(x,x')} > -\infty.$$

For every integer k, let A_k be the set of $x \in M \setminus \mathcal{E}^M$ such that

$$\psi(x') \geq \psi(x) - k\, d^M(x,x') \quad \forall x' \in B^M(x, 1/k), \tag{4.4}$$

where d^M denotes a geodesic distance on M coming from a fixed Riemannian metric on M and B^M stand for the open balls with respect to d^M. Then if we denote by $\bar{A}_k$ a Borel set of full Lebesgue measure consisting only of Lebesgue density points of each A_k, we have $M \setminus (\mathcal{E}^M \cup \mathcal{E}') = \cup_k \bar{A}_k$ for some Borel set $\mathcal{E}'$ of Lebesgue measure zero. For each k, there is a finite set of points $z_1^k, \ldots, z_{m_k}^k$ in $\bar{A}_k$ such that

$$\bar{A}_k \subset \bigcup_{l=1}^{m_k} B^M(z_l^k, 1/2k).$$

For every $k \in \mathbb{N}$ and every $l \in \{1, \ldots, m_k\}$ we define the function $\psi_{k,l} : M \to \mathbb{R}$ by

$$\psi_{k,l}(x) := \sup\Big\{\psi(z) - k\, d^M(x,z) \,|\, z \in \bar{A}_k \cap B^M(z_l^k, 1/2k)\Big\} \qquad \forall x \in M.$$

By construction and (4.4), $\psi_{k,l}$ coincides with ψ on $\bar{A}_k \cap B^M(z_l^k, 1/2k)$ and in addition $\psi_{k,l}$ is k-Lipschitz on M. Therefore by Rademacher's theorem, $\psi_{k,l}$ is differentiable almost everywhere on M. Then for each $k \in \mathbb{N}$ and each $l \in \{1, \ldots, m_k\}$, there is a Borel set $B_{k,l} \subset B^M(z_l^k, 1/2k)$ of Lebesgue measure zero such that for every $x \in \bar{A}_k \setminus B_{k,l}$, ψ is approximately differentiable at x. The proof of the almost everywhere differentiability of ϕ follows the same lines. $\square$

Returning to the proof of Theorem 1.4, we apply Lemma 4.2 and set

$$\mathcal{N}' := \mathcal{N} \cup \left(\mathcal{F}^M \times N\right) \cup \left(M \cup \mathcal{F}^N\right) \quad \text{and} \quad \mathcal{S}' := \mathcal{S} \setminus \mathcal{N}'.$$

By construction, if $((x_1, y_1), \ldots, (x_L, y_L))$ is a $(0, \mathcal{S}')$-chain then ψ is approximately differentiable at all the x_l with $l = 1, \ldots, L$ and ϕ is approximately differentiable at all the y_l with $l = 1, \ldots, L$. On the other hand, by (4.1)-(4.3), we have for every $l = 1, \ldots, L-1$

$$y_l \neq y_{l+1} \Longrightarrow \left\{ \begin{array}{rcll} c(x', y_l) & \geq & \phi(y_l) - \psi(x') & \forall x' \in M \quad \text{(with equality at } x_l) \\ c(x', y_{l+1}) & \geq & \phi(y_{l+1}) - \psi(x') & \forall x' \in M \quad \text{(with equality at } x_l) \end{array} \right.$$

and

$$x_l \neq x_{l+1} \Longrightarrow \left\{ \begin{array}{rcll} c(x_l, y') & \geq & \phi(y') - \psi(x_l) & \forall y' \in N \quad \text{(with equality at } y_l) \\ c(x_{l+1}, y') & \geq & \phi(y') - \psi(x_{l+1}) & \forall y' \in N \quad \text{(with equality at } y_l) \end{array} \right.$$

Thus this shows that for every $l = 1, \ldots, L-1$, either $y_l \neq y_{l+1}$ and both functions $x' \mapsto c(x', y_l)$ and $x' \mapsto c(x', y_{l+1})$ have $-d_{x_l}\psi$ as a common approximate sub-differential at x_l, or $x_l \neq x_{l+1}$ and both functions $y' \mapsto c(x_l, y')$ and $y' \mapsto c(x_{l+1}, y')$ have $d_{y_l}\phi$ as a common approximate sub-differential at y_l. In other words, any $(0, \mathcal{S}')$-chain is a $(c, \mathcal{S}')$-chain. The second part of Theorem 1.4 follows by Theorem 2.6.

4.2 Proof of Theorem 1.5

Let $c : M \times N \to \mathbb{R}$ be a continuous cost and $\epsilon > 0$ be fixed, without loss of generality smoothing c if necessary we may assume that c is smooth. First, we consider cubic subdivisions of M and N. There are two finite families of cubes $\{\mathcal{I}_k^M\}_{k \in K}$ and $\{\mathcal{I}_l^N\}_{l \in L}$ with disjoint interiors such that (see Figure 3)

$$M = \bigcup_{k \in K} \mathcal{I}_k^M \quad \text{and} \quad N = \bigcup_{l \in L} \mathcal{I}_l^N.$$

This means that for every $k \in K$ (resp. $l \in L$) there is a smooth diffeomorphism $\Phi_k^M : \mathcal{V}_k^M \to \mathbb{R}^n$ (resp. $\Phi_l^N : \mathcal{V}_l^N \to \mathbb{R}^n$) defined on an open neighborhood $\mathcal{V}_k^M$ of $\mathcal{I}_k^M$ (resp. $\mathcal{V}_l^N$ of $\mathcal{I}_l^N$) such that

$$\Phi_k^M\left(\mathcal{I}_k^M\right) = I^n := [0,1]^n \quad (\text{resp. } \Phi_l^N\left(\mathcal{I}_l^N\right) = I^n).$$

For every $k \in K$ and $l \in L$, we set $\mathcal{P}_{k,l} := \mathcal{I}_k^M \times \mathcal{I}_l^N$ and we define the map $\Phi_{k,l} : \mathcal{P}_{k,l} \to P := I^n \times I^n$ by $\Phi_{k,l}(x, y) = (\Phi_k^M(x), \Phi_l^N(y))$.

Lemma 4.3. *There are $C > 0$ and two families of Lipschitz continuous functions $\{f_{k,l}\}_{k \in K, l \in L}, \{h_{k,l}\}_{k \in K, l \in L} : P \to \mathbb{R}$ which satisfy the following properties for every $k \in K$, $l \in L$ (we set $f := f_{k,l}$, $h := h_{k,l}$, $\Phi := \Phi_{k,l}$ and $\mathcal{P} := \mathcal{P}_{k,l}$)*

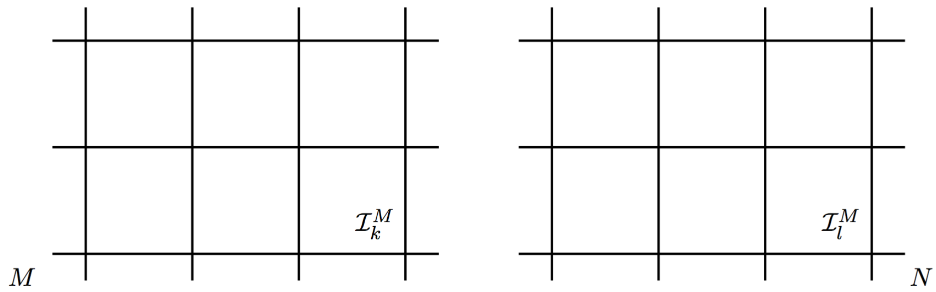

Figure 3: Cubic subdivisions of M and N

(i) *For every* (x, y) *in* $\partial\mathcal{P}$, $h(\Phi(x, y)) = c(x, y)$.

(ii) *For every* $(u, v) \in P$, *we have* $h(u, v) = f(u, v) + C|u|^2 + C|v|^2$.

(iii) *The function* f *is concave and locally piecewise affine in the interior of* P, *that is, there is a locally finite partition* $\{K^i = K^i_{k,l}\}_{i\in\mathbb{N}}$ *of Int*(P), *with each* K^i *a compact convex subset of Int*(P), *such that* f *is affine on each set of the partition. Moreover, for every* (u, v) *in the boundary of some* K^i, *one of the approximate sub-differentials* $\tilde{D}^-_u f$ *or* $\tilde{D}^-_v f$ *is empty at* (u, v).

(iv) $\|c - h \circ \Phi\|_{C^0} < \epsilon/3$.

Proof of Lemma 4.3. Fix $k \in K$ and $l \in L$, since c is smooth in a neighborhood of $\mathcal{P}$ there is a constant $C > 0$ such that the smooth function $\bar{f} : P \to \mathbb{R}$ defined by

$$\bar{f}(u, v) := \left(c \circ \Phi^{-1}\right)(u, v) - C|u|^2 - C|v|^2 \qquad \forall (u, v) \in P,$$

is uniformly concave. We note that since the set of sets of the form $\mathcal{P} = \mathcal{P}_{k,l} := \mathcal{I}_k^M \times \mathcal{I}_l^N$ is finite, we can take the same constant C for all sets $\mathcal{P}$. Let $d(\cdot, \partial P)$ denote the distance function to the boundary of P in $\mathbb{R}^n \times \mathbb{R}^n$. By uniform concavity, for every $(u, v) \in \mathrm{Int}(P)$ there are $p, q \in \mathbb{R}^n$ and $\rho \in (0, d((u, v), \partial P))$ such that

$$\bar{f}(u', v') \leq \bar{f}(u, v) + p \cdot (u' - u) + q \cdot (v' - v) \qquad \forall (u', v') \in P \tag{4.5}$$

and

$$\bar{f}(u', v') \geq$$
$$\bar{f}(u, v) + p \cdot (u' - u) + q \cdot (v' - v) - \frac{\epsilon d((u, v), \partial P))}{6} \qquad \forall (u', v') \in B\left((u, v), \rho\right). \tag{4.6}$$

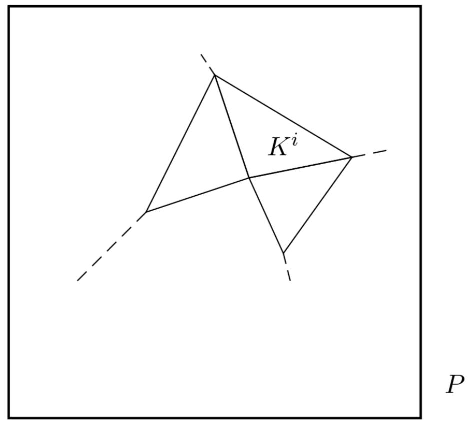

Figure 4: Locally finite partition of $\mathrm{Int}(P)$ by convex sets as in Lemma 4.3 (iii)

Then by local compactness, there is a locally finite family $\{(u_i, v_i)\}_{i\in\mathbb{N}}$ in $\mathrm{Int}(P)$ associated with families $\{(p_i, q_i)\}_{i\in\mathbb{N}}, \{\rho_i\}_{i\in\mathbb{N}}$ such that

$$P \subset \bigcup_{i\in\mathbb{N}} B\left((u_i, v_i), \rho_i\right) \quad \text{and} \quad \nabla \bar{f}(u_i, v_i) = (p_i, q_i)\, \forall i \in \mathbb{N}.$$

Define the function $f : P \to \mathbb{R}$ by

$$f(u,v) := \begin{cases} \min_{i\in\mathbb{N}}\{\bar{f}(u_i, v_i) + p_i \cdot (u - u_i) + q_i \cdot (v - v_i)\} & \text{if} \quad (u,v) \in \mathrm{Int}(P) \\ \bar{f}(u,v) & \text{if} \quad (u,v) \in \partial P. \end{cases}$$

By construction, f is concave, Lipschitz continuous on P, locally piecewise affine, it coincides with $\bar{f}$ on ∂P and in addition by (4.5)-(4.6) it satisfies $\bar{f} \leq f \leq \bar{f} + \epsilon/6$. Moreover, by uniform concavity, if we denote by $\{K^i\}_{i\in\mathbb{N}}$ the maximal partition such that f is affine on each K^i, then the differentials of f in $\mathrm{Int}(K^i)$ and $\mathrm{Int}(K^j)$ with $i \neq j$ are different. So that the sub-differential of f in both variable on the boundary of each K^i is empty, which implies the desired property in (iii). In conclusion, setting $h(u,v) := f(u,v) + C|u|^2 + C|v|^2$, we get the result. $\square$

Applying Lemma 4.3 to all products $\mathcal{P}_{k,l}$ with $k \in K$ and $l \in L$, we obtain a Lipschitz continuous function $\bar{c} : M \times N \to \mathbb{R}$ such that $\|\bar{c} - c\|_{C^0} < \epsilon/3$, along with a countable family of compact sets $\{K^i_{k,l}\}_{i\in\mathbb{N},k\in K,l\in L}$ such that each $c \circ (\Phi_{k,l})^{-1}$ has the

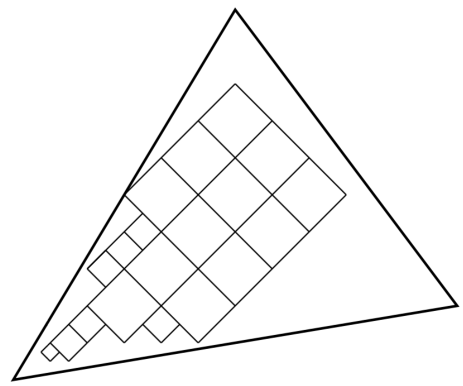

Figure 5: A countable tiliing of K by l^1-balls

form given in Lemma 4.3 (ii)-(iii) on P.

Let $\|\cdot\|$ be the l^1-norm in $\mathbb{R}^n \times \mathbb{R}^n$, that is the norm defined by

$$\|(u,v)\| = |u_1| + \ldots + |u_n| + |v_1| + \ldots |v_n| \qquad \forall (u,v) \in \mathbb{R}^n \times \mathbb{R}^n.$$

We denote by $B^{\|\cdot\|}((u,v),r)$ the closed ball centered at (u,v) with radius $r \geq 0$ with respect to $\|\cdot\|$. Given a l^1-ball $B = B^{\|\cdot\|}((u,v),r) \subset \mathbb{R}^n \times \mathbb{R}^n$ and $N \in \mathbb{N}$ with $N \geq 2$, we call N-subpartition of B the partition of B into N^{2n} l^1-balls of the same radius, that is the unique covering of B of the form

$$B = \bigcup_{t=1}^{N^{2n}} B^{\|\cdot\|}((u_t,v_t),r') \quad \text{with} \quad r' = \frac{r}{N},$$

for some $N \in \mathbb{N}$ such that all the sets $B^{\|\cdot\|}((u_t,v_t),r')$ have disjoint interior. The following result is illustrated in Figure 5.

Lemma 4.4. *Let K be a compact convex set in $\mathbb{R}^{2n} = \mathbb{R}^n \times \mathbb{R}^n$ with nonempty interior, then there are a family $\{w_j\}_{j\in\mathbb{N}}$ in Int(K) and a family of positive real numbers $\{\rho_j\}_{j\in\mathbb{N}}$ such that*

$$\mathit{Int}(K) = \bigcup_{j\in\mathbb{N}} B^{\|\cdot\|}(w_j,\rho_j)$$

and
$$Int\Big(B^{\|\cdot\|}(w_j,\rho_j)\Big)\cap Int\Big(B^{\|\cdot\|}(w_{j'}),\rho_{j'})\Big)=\emptyset \qquad \forall j\neq j'\in\mathbb{N},$$
Such a covering of Int(K) is called a countable tiling of K by l^1-balls.

Proof of Lemma 4.4. Since K is compact with nonempty interior, there are $w\in K$ and $k>k'$ with $k\in\mathbb{N}, k'\in\mathbb{Z}$ such that

$$B^{\|\cdot\|}(w,2^{-k})\subset \mathrm{Int}(K)\subset K\subset B^{\|\cdot\|}(w,2^{-k'}).$$

By construction, the ball $B^{\|\cdot\|}(w,2^{-k})$ contains one of the small l^1-balls in an N-subpartition of $B:=B^{\|\cdot\|}(w,2^{-k'})$, of the form

$$B=\bigcup_{t=1}^{N^{2n}}B^{\|\cdot\|}(w_t,r),$$

with $N>2^{k-k'}$ and $r:=2^{-k'}/N$. Let J^0 be the set of indices $t\in\{1,\ldots,N^{2n}\}$ such that $B^{\|\cdot\|}(w_t,r)\subset\mathrm{Int}(K)$ and K^0 defined as

$$K_0:=\bigcup_{t\in J^0}B^{\|\cdot\|}(w_t,r)\subset\mathrm{Int}(K).$$

We consider now a 2-subpartition of the above subpartition, that is to say a $2N$-subpartition of B of the form

$$B=\bigcup_{t=1}^{N^1}B^{\|\cdot\|}(w_t^1,r^1),$$

with $N^1:=2N$ and $r^1:=r/2$. By construction, each $B^{\|\cdot\|}(w_t,r)$ corresponds to a block of balls in the N^1-subpartition. Then we consider the set J^1 of indices $t\in\{1,\ldots,(N^1)^{2n}\}$ such that $B^{\|\cdot\|}(w_t^1,r^1)\subset\mathrm{Int}(K)$ and $B^{\|\cdot\|}(w_t^1,r^1)\cap\mathrm{Int}(K^0)=\emptyset$. By construction, the balls $B^{\|\cdot\|}(w_t,r)$ with $t\in J^0$ and $B^{\|\cdot\|}(w_{t'}^1,r^1)$ with $t'\in J^1$ have disjoint interiors, the balls $B^{\|\cdot\|}(w_t^1,r^1)$, $B^{\|\cdot\|}(w_{t'}^1,r^1)$ with $t\neq t'\in J^1$ have disjoint interiors and the set

$$K^1:=K^0\cup\bigcup_{t\in J^1}B^{\|\cdot\|}\left(w_t^1,r^1\right)$$

is contained $\mathrm{Int}(K)$. We conclude the construction by considering new subpartitions of the initial N-subpartition of B. □

From Lemma 4.4, each set $K_{k,l}^i \subset P$ given by Lemma 4.3 (with $k \in K, l \in L, i \in \mathbb{N}$) admits a countable tiling by l^1-balls of the form

$$K_{k,l}^i = \bigcup_{j\in\mathbb{N}} B^{\|\cdot\|}\left(w_j^{k,l,i}, \rho_j^{k,l,i}\right).$$

Relabelling the family of sets $(\Phi_{k,l})^{-1}(B^{\|\cdot\|}(w_j^{k,l,i}, \rho_j^{k,l,i}))$ with $k \in K, l \in L, i, j \in \mathbb{N}$, we get a family of closed sets $\{\mathcal{C}_s\}_{s\in\mathbb{N}}$ in $M \times N$. In the sequel, we denote by $k(s), l(s)$ the indices such that $\mathcal{C}_s$ is contained in $\mathcal{I}_k^M \times \mathcal{I}_l^N$ and we set

$$\Phi_s := \Phi_{k(s),l(s)} \quad \text{and} \quad C_s := \Phi_s\left(\mathcal{C}_s\right) \subset P \qquad \forall s \in \mathbb{N}.$$

By construction, for each $s \in \mathbb{N}$, the function $\bar{c} \circ (\Phi_s)^{-1}$ is the sum of an affine function and of the function $C|u|^2 + C|v|^2$ on C_s.

The proof of the following lemma is left to the reader.

Lemma 4.5. *For every $r' > 0$, the function (we set $B^{\|\cdot\|}(r') := B\|\cdot\|((0,0), r')$)*

$$D = d(\cdot, \partial B^{\|\cdot\|}(r')) : B\|\cdot\|((0,0), r') \to [0, r']$$

given by the distance function to $\partial B^{\|\cdot\|}(r')$ (w.r.t. the Euclidean norm) is concave and piecewise affine. Moreover it is differentiable on the set Diff(D) of (u,v) in $B^{\|\cdot\|}(r')$ with $u_i, v_i \neq 0$ for all $i = 1, \ldots, n$ and satisfies

$$\left|\frac{\partial D}{\partial u_i}(u,v)\right| = \left|\frac{\partial D}{\partial v_i}(u,v)\right| = \frac{1}{\sqrt{2n}} \qquad \forall (u,v) \in \mathit{Diff}(D).$$

Given a N-subpartition (with $N \in \mathbb{N}$) of some l^1-ball $B = B^{\|\cdot\|}((u,v), r)$ of the form

$$B = \bigcup_{t=1}^{N^{2n}} B^{\|\cdot\|}((u_t, v_t), r') \quad \text{with} \quad r' = \frac{r}{N}$$

and $\delta > 0$, we define the function $D^\delta : B \to \mathbb{R}$ by

$$D^\delta(u,v) = -\delta\, d\left((u,v), B^{\|\cdot\|}(u_t, v_t), r'\right) \quad \text{if} \quad (u,v) \in B^{\|\cdot\|}((u_t, v_t), r').$$

By construction, D^δ is Lipschitz continuous, piecewise affine, it vanishes on the boundary of each $B^{\|\cdot\|}((u_t, v_t), r')$ (in particular it vanishes on ∂B), it is convex on each l^1-ball $B^{\|\cdot\|}((u_t, v_t), r')$, it is differentiable on an open set $\mathrm{Diff}(D^\delta)$ which is the complement of a finite union of horizontal and vertical hyperplanes, and it satisfies

$$\left|\frac{\partial D^\delta}{\partial u_i}(u,v)\right| = \left|\frac{\partial D^\delta}{\partial v_i}(u,v)\right| = \frac{\delta}{\sqrt{2n}} \qquad \forall (u,v) \in \mathrm{Diff}(D) \cap B^{\|\cdot\|}((u_t, v_t), r'),\ \forall t, \quad (4.7)$$

for every $i = 1, \ldots, n$. Moreover, we have

$$0 \geq D^{\delta}(u,v) \geq -\frac{r\delta}{N} \qquad \forall (u,v) \in B. \tag{4.8}$$

We need now to associate subpartitions to each set $\mathcal{C}_s$, that is subpartitions of C_s via Φ_s. Before stating the result we consider g^M, g^N two smooth Riemannian metrics on M and N respectively, we denote by $|\cdot|^M, |\cdot|^N$ the corresponding norms and by ∇^M, ∇^N the corresponding connections. Then given a function $a : M \times N \to \mathbb{R}$ differentiable at some (x,y), we denote by $\nabla_x^M a(x,y)$ and $\nabla_y^N a(x,y)$ the gradients of a with respect to x and y.

Lemma 4.6. *There are a sequence $\{N_s\}_{s\in\mathbb{N}}$ of integers greater than one and two sequences $\{\delta_s\}_{s\in\mathbb{N}}, \{\kappa_s\}_{s\in\mathbb{N}}$ of positive real numbers such that the continuous function $\tilde{c} : M \times N \to \mathbb{R}$ defined by*

$$\tilde{c}(x,y) = \bar{c}(x,y) + D^{\delta_s}\left(\Phi_s(x,y)\right) \qquad \forall (x,y) \in \mathcal{C}_s,\ \forall s \in \mathbb{N}, \tag{4.9}$$

(where $D^{\delta_s} : C_s \to \mathbb{R}$ denotes the function associated with the N_s-subpartition of C_s as above) satisfies the following properties:

(i) $\|\tilde{c} - \bar{c}\|_{C^0} < \epsilon/3$.

(ii) For every $s \in \mathbb{N}$ and every $(x,y) \in \mathcal{C}_s$ whose image by Φ_s belongs to the boundary of a cube of the N_s-subpartition, both sub-differentials $\tilde{D}_x^- \tilde{c}(x,y)$ and $\tilde{D}_y^- \tilde{c}(x,y)$ are empty.

(iii) There are sets of Lebesgue measure zero $M_0 \subset M$ and $N_0 \subset N$ such that for every $(x,y) \in M \times N \setminus ((M_0 \times N) \cup (M \times N_0))$, either both sub-differentials $\tilde{D}_x^- \tilde{c}(x,y)$ and $\tilde{D}_y^- \tilde{c}(x,y)$ are empty, or $\tilde{c}$ is differentiable at (x,y).

(iv) For every $(x,y) \in \mathcal{C}_s$ such that $\tilde{c}$ is differentiable at (x,y) and every $(x,y') \in \mathcal{C}_{s'}$ such that $\tilde{c}$ is differentiable at (x,y'), there holds

$$\left|\nabla_x^M \tilde{c}(x,y) - \nabla_x^M \tilde{c}(x,y')\right|^M < \kappa_s + \kappa_{s'} \implies s = s'.$$

(v) For every $(x,y) \in \mathcal{C}_s$ such that $\tilde{c}$ is differentiable at (x,y) and every $(x',y) \in \mathcal{C}_{s'}$ such that $\tilde{c}$ is differentiable at (x',y), there holds

$$\left|\nabla_y^N \tilde{c}(x,y) - \nabla_y^N \tilde{c}(x',y)\right|^N < \kappa_s + \kappa_{s'} \implies s = s'.$$

Proof of Lemma 4.6. Since the maps Φ_k^M, Φ_l^N with $k \in K, l \in L$ are smooth diffeomorphism in open neighborhoods of $\mathcal{I}_k^M$ and $\mathcal{I}_l^N$ respectively, there is a constant $\lambda \in (0,1)$

such that for every cost $\tilde{c}: M \times N \to \mathbb{R}$ of the form (4.9), thanks to the formula (4.7), we have for each point (x, y) of differentiability of $\tilde{c}$ in $\mathcal{C}_s$,

$$\nabla_x^M \tilde{c}(x,y) = \nabla_x^M \bar{c}(x,y) + V \quad \text{and} \quad \nabla_y^N \tilde{c}(x,y) = \nabla_y^N \bar{c}(x,y) + W \tag{4.10}$$

with

$$\lambda\, \delta_s \leq |V|^M \leq \frac{\delta_s}{\lambda} \quad \text{and} \quad \lambda\, \delta_s \leq |W|^N \leq \frac{\delta_s}{\lambda}. \tag{4.11}$$

Let $L(\bar{c}) > 0$ be the Lipschitz constant of $\bar{c}$, we define the two sequences $\{\mu_s\}_{s\in\mathbb{N}}, \{\nu_s\}_{s\in\mathbb{N}}$ recursively by

$$\mu_0 = 0, \quad \nu_0 = 1, \quad \mu_{s+1} := \frac{\nu_s}{\lambda^2} + \frac{3L(\bar{c})}{\lambda}, \quad \nu_{s+1} := \mu_{s+1} + 1 \qquad \forall s \in \mathbb{N},$$

then we set

$$\delta_s := \frac{\mu_s + \nu_s}{2} \quad \text{and} \quad \kappa_s := \lambda \left(\frac{\nu_s - \mu_s}{2} \right) \qquad \forall s \in \mathbb{N}.$$

By (4.10)-(4.11), we check easily that for every point (x, y) of differentiability of $\tilde{c}$ in $\mathcal{C}_s$,

$$\left|\nabla_x^M \tilde{c}(x,y)\right|^M - \kappa_s \geq \lambda \delta_s - L(\bar{c}) - \kappa_s = \lambda\, \mu_s - L(\bar{c})$$

and

$$\begin{aligned} \left|\nabla_x^M \tilde{c}(x,y)\right|^M + \kappa_s &\leq \frac{\delta_s}{\lambda} + L(\bar{c}) + \kappa_s \\ &\leq \frac{\delta_s}{\lambda} + L(\bar{c}) + \frac{\nu_s - \mu_s}{2\lambda} = \frac{\nu_s}{\lambda} + L(\bar{c}) < \lambda\, \mu_{s+1} - L(\bar{c}). \end{aligned}$$

This shows that (iv) is satisfied. We can repeat the same argument to prove (v). To fulfill property (i), thanks to (4.8) we just need to take each N_s large enough. Finally the property (ii)-(iii) hold by construction. $\square$

By the previous lemma, if $\kappa : M \times N \to \mathbb{R}$ is a Lipschitz function such that

$$\left|\nabla_x^- \kappa(x,y)\right|^M, \left|\nabla_y^- \kappa(x,y)\right|^N < \kappa_s \qquad \forall (x,y) \in \mathcal{C}_s, \forall s \in \mathbb{N}, \tag{4.12}$$

(where $|\nabla_x^- \kappa|$ (resp. $|\nabla_y^- \kappa|$) denotes the maximum of the norm of the gradients of support functions of class C^1 that can be put under the graph of $\kappa(\cdot, y)$ at x (resp. of $\kappa(x, \cdot)$ at y)) and

$$\left\{ \begin{array}{lcl} \tilde{D}_x^- \tilde{c}(x,y) = \emptyset & \Longrightarrow & \tilde{D}_x^- \hat{c}(x,y) = \emptyset \\ \tilde{D}_y^- \tilde{c}(x,y) = \emptyset & \Longrightarrow & \tilde{D}_x^- \hat{c}(x,y) = \emptyset \end{array} \right. \qquad \forall (x,y) \in \mathcal{C}_s, \forall s \in \mathbb{N}, \tag{4.13}$$

where the cost $\hat{c} : M \times N \to \mathbb{R}$ is given by

$$\hat{c}(x,y) := \tilde{c}(x,y) + \kappa(x,y) \qquad \forall (x,y) \in M \times N,$$

then any eventual $(\hat{c}, \mathcal{S})$-chain in $M \times N \setminus ((M_0 \times N) \cup (M \times N_0))$ (with $\mathcal{S}$ an eventual $(\hat{c}, \mu, \nu)$-minimizing set) would be confined in a set $\mathcal{C}_s$ for some $s \in \mathbb{N}$. As a matter of fact, if $(x,y), (x',y') \in M \times N \setminus ((M_0 \times N) \cup (M \times N_0))$ with $x = x'$ (the case $y = y'$ is left to the reader), $(x,y) \in \mathcal{C}_s$, $(x,y') \in \mathcal{C}_{s'}$ satisfy

$$\tilde{D}_x^- \hat{c}(x,y) \cap \tilde{D}_x^- \hat{c}(x,y) \neq \emptyset,$$

then by Lemma 4.6 (iii) and (4.12)-(4.13), the function $\tilde{c}$ is differentiable at (x,y) and (x,y') and there are $p, q \in T_x M$ such that

$$\nabla_x^M \tilde{c}(x,y) + p = \nabla_x^M \tilde{c}(x,y') + q \quad \text{and} \quad |p|^M < \kappa_s, \, |q|^M < \kappa_{s'},$$

and so by Lemma 4.6 (iv) there holds $s = s'$. Fix $s \in \mathbb{N}$ and define the function $h_s : C_s \to \mathbb{R}$ by

$$\begin{aligned} h_s(u,v) &:= \left(\hat{c} \circ (\Phi_s)^{-1}\right)(u,v)) && (4.14) \\ &= f_s(u,v) + C|u|^2 + C|v|^2 + D^{\delta_s}(u,v) + k_s(u,v) && (4.15) \end{aligned}$$

for every $(u,v) \in C_s$, where f_s is affine and $k_s := \kappa \circ (\Phi_s)^{-1}$. There is a one-to-one correspondence between the $\hat{c}$-chains in $\mathcal{C}_s$ and the h_s-chains in C_s.

Lemma 4.7. *Let*

$$\left\{(x_1,y_1), \ldots (x_L,y_L)\right\} \in \mathcal{C}_s$$

with $L \geq 1$ and

$$(u_i,v_i) := \Phi_s(x_i,y_i) = \Phi_{k(s),l(s)}(x_i,y_i) = \left(\Phi_{k(s)}^M(x_i), \Phi_{l(s)}^N(y_i)\right) \in C_s \qquad \forall l = 1, \ldots, L.$$

Then $((x_1,y_1), \ldots (x_L,y_L))$ is a $\hat{c}$-chain if and only if $((u_1,v_1), \ldots (u_L,u_L))$ is a h_s-chain.

Proof of Lemma 4.7. Let $\{(x_1,y_1), \ldots (x_L,y_L)\} \in \mathcal{C}_s$ be fixed. By construction, for every l, $\hat{c}(x_l,y_l) = h_s(u_l,v_l)$. Hence the set $\{(x_1,y_1), \ldots (x_L,y_L)\}$ is $\hat{c}$-cyclically monotone if and only if the set $\{(u_1,v_1), \ldots (u_L,v_L)\}$ is $\hat{c}$-cyclically monotone. By construction,

$$\hat{c} = h_s \circ \Phi_s = h_s \circ \left(\Phi_{k(s)}^M, \Phi_{l(s)}^N\right) \quad \text{on} \quad \mathcal{C}_s$$

and $\Phi_{k(s)}^M : \mathcal{I}_{k(s)}^M \to I^n$, $\Phi_{l(s)}^N : \mathcal{I}_{l(s)}^N \to I^n$ are smooth diffeomorphisms. So we have formally

$$\tilde{D}_x^- \hat{c}(x,y) = \tilde{D}^-_{\Phi_{k(s)}^M(x)} h_s(x,y) \circ D_x \Phi_{k(s)}^M \quad \text{and} \quad \tilde{D}_y^- \hat{c}(x,y) = \tilde{D}^-_{\Phi_{l(s)}^N(x)} h_s(x,y) \circ D_y \Phi_{l(s)}^N.$$

We conclude easily. $\square$

As a consequence of the above discussion, a way to approach c by a uniquely minimizing cost c^ϵ such that $\|c^\epsilon - c\|_{C^0} < \epsilon$ is to construct a Lipschitz function $\kappa : M \times N \to \mathbb{R}$ satisfying (4.12)-(4.13) and such that each cost h_s given by (4.14) is wandering in C_s.

Let s be fixed. By construction, C_s is a l^1-ball of radius $r > 0$ equipped with a N_s-subpartition of the form

$$C_s = \bigcup_{t=1}^{N_s^{2n}} B^{\|\cdot\|}\left(w_t, r/N_s\right)$$

and the function D^{δ_s} is adapted to the subpartition. Set $T := N_s^{2n}$ and $r' := r/N_s$, and given a T-tuple $\tilde{\delta} = (\tilde{\delta}_1, \ldots, \tilde{\delta}_t)$ of positive real numbers define the function $\tilde{D}^{\tilde{\delta}} : C_s \to \mathbb{R}$ by

$$\tilde{D}^{\tilde{\delta}}(w) = -\tilde{\delta}_t\, d\left(w, B^{\|\cdot\|}(w_t, r')\right) \quad \text{if} \quad w \in B^{\|\cdot\|}(w_t, r').$$

From Lemmas 4.5 and 4.6, if the real numbers $\tilde{\delta}_1, \ldots, \tilde{\delta}_t$ are distinct and sufficiently close to δ_s, then there is $\kappa > 0$ such that any eventual chain for the cost

$$(u, v) \in C_s \longmapsto f(u, v) + C|u|^2 + C|v|^2 + \tilde{D}^{\tilde{\delta}}(u, v) + k(u, v),$$

with $k : C_s \to \mathbb{R}$ Lipschitz such that $|\nabla_u^- k|, |\nabla_v^- k| < \kappa$, is confined in a l^1-ball of the form $B^{\|\cdot\|}(w_t, r')$. So modifying a little the construction of $\hat{c}$, we may assume that the above property holds and we are left to show how to construct a wandering cost on a l^1-ball of the form $B^{\|\cdot\|}(w_t, r')$.

From now on, s and t are fixed and we set

$$B := B^{\|\cdot\|}(w_t, r') \subset C_s, \quad D := -\tilde{\delta}_t\, d\left(w, B\right)$$

$$\begin{aligned} \text{and} \quad h(u, v) &:= f(u, v) + C|u|^2 + C|v|^2 + D(u, v) \\ &= p \cdot u + q \cdot v + C|u|^2 + C|v|^2 + D(u, v) \qquad \forall (u, v) \in C_s. \end{aligned}$$

In order to conclude the proof of Theorem 1.5, our objective is now to construct a Lipschitz function $k : B \to \mathbb{R}$ which satisfies the following properties:

(P1) $k = 0$ on ∂B.

(P2) $-\epsilon/3 \le k \le \min\{\epsilon/3, -D(u, v)/2\}$ on B.

(P3) The cost $h + k$ is wandering on $\mathrm{Int}(B)$.

We shall cover B with a countable tiling of l^∞-balls. Set $Q := [-1/2, 1/2]^n \times [-1/2, 1/2]^n$ and for every $\lambda > 0$, $\lambda Q := [-\lambda/2, \lambda/2]^n \times [-\lambda/2, \lambda/2]^n$.

Lemma 4.8. *There are $L_2 > L_1 > 0$ such that for every $\lambda > 0$, there exist a $(\mathcal{L}^n, \mathcal{L}^n)$-null set $\mathcal{N}^\lambda \subset \lambda Q$ containing $\partial(\lambda Q)$ and a Lipschitz function $\sigma^\lambda : \lambda Q \to \mathbb{R}$ which satisfy the following properties:*

(i) $\sigma^\lambda = 0$ *on* $\partial(\lambda Q)$.

(ii) For every (u, v) *in* $Int(\lambda Q) \setminus \mathcal{N}^\lambda$, *any* $p \in \tilde{D}_u^- \sigma^\lambda(u, v)$ *and* $q \in \tilde{D}_v^- \sigma^\lambda(u, v)$ *satisfy* $|p|, |q| \in [L_1/\lambda, L_2/\lambda]$. *In particular,* σ^λ *is* (L_2/λ)*-Lipschitz.*

(iii) There are no $(\sigma^\lambda, \lambda Q \setminus \mathcal{N}^\lambda)$*-chains of length* 4.

Proof of Lemma 4.8. Let $S^n \subset \mathbb{R}^{n+1} = \mathbb{R}^n \times \mathbb{R}$ be the sphere of radius 1 with center on the vertical axis and passing through the origin 0. Then consider a smooth function $F : Q \to S^n$ such that F is a diffeomorphism from $\mathrm{Int}(Q)$ to $S^n \setminus \{0\}$ and $F(\partial Q) = 0$. Then define the cost $\bar{\sigma} : Q \to \mathbb{R}$ by

$$\bar{\sigma}(u, v) := F(u) \cdot F(v) \qquad \forall (u, v) \in Q.$$

By construction, $\bar{\sigma} = 0$ on ∂Q and there are no $(\mathcal{S}, \bar{\sigma})$-chains of length 4 with $\mathcal{S} = \mathrm{Int}(Q)$. As a matter of fact, if

$$\Big((u_1, v_1), (u_2, v_2), (u_3, v_3), (u_4, v_4)\Big) \in \mathcal{S}^4$$

is a chain satisfying $v_1 = v_2, v_2 \neq v_3, v_3 = v_4$ (the case $u_1 = u_2$ is left to the reader), then we have

$$v_2 \neq v_3 \quad \text{and} \quad F(v_2) - F(v_3) = \lambda \vec{n},$$

with $\lambda \in \mathbb{R} \setminus \{0\}$ and $\vec{n}$ is the outward normal to S^n at $F(u_2) = F(u_3)$. By $\bar{\sigma}$-cyclical monotonicity, we have

$$\begin{cases} (F(v_2) - F(v_4)) \cdot (F(u_2) - F(u_4)) & \leq & 0 \\ (F(v_1) - F(v_3)) \cdot (F(u_1) - F(u_3)) & \leq & 0, \end{cases}$$

which implies that the two quantities

$$(F(u_2) - F(u_4)) \cdot \vec{n} \quad \text{and} \quad (F(u_1) - F(u_2)) \cdot \vec{n}$$

have the same sign. This contradicts the uniform convexity of $\mathbb{S}^n$ and the fact that $F(u_2)$, $F(u_1) \neq F(u_2)$ and $F(u_4) \neq F(u_2)$ belong to $\mathbb{S}^n$.

Consider a countable tiling of $\mathrm{Int}(Q)$ by l^1-balls of the form

$$\mathrm{Int}(Q) = \bigcup_{j \in \mathbb{N}} B^{\|\cdot\|}(w_j, \rho_j)$$

and for every $A>0$, define the function $\mathcal{D}^A : Q \to \mathbb{R}$ by

$$\mathcal{D}^A(u,v) = -A\, d\left(w, B^{\|\cdot\|}(w_j,\rho_j)\right) \quad \text{if} \quad w \in B^{\|\cdot\|}(w_j,\rho_j).$$

Denote by Q_0 the union over $j \in \mathbb{N}$ of all sets of the form

$$\left\{ w_j + z \,|\, z = (z_1,\ldots,z_{2n}) \text{ s.t. } z_i = 0 \text{ for some } i \right\} \cap B^{\|\cdot\|}(w_j,\rho_j),$$

by construction, it is a $(\mathcal{L}^n,\mathcal{L}^n)$-null set in $\mathrm{Int}(Q)$. Denote by $\mathcal{B}$ the unions of the boundaries of all l^1-balls $B^{\|\cdot\|}(w_j,\rho_j)$. By construction, both sub-differentials $\tilde{D}_u^-\mathcal{D}^A$ and $\tilde{D}_v^-\mathcal{D}^A$ are empty over $\mathcal{B}$. The remaining set $\mathcal{Q} := \mathrm{Int}(Q) \setminus (Q_0 \cup \mathcal{B})$ is the union of countably many convex domains where $\mathcal{D}^A$ is affine with slope $\nabla_u\mathcal{D}^A, \nabla_v\mathcal{D}^A$ among the values (see Lemma 4.5)

$$\frac{A}{\sqrt{2n}}E \quad \text{with} \quad E = (E_1,\ldots,E_n) \quad \text{and} \quad E_1,\ldots,E_n \in \{\pm 1\}.$$

For every $A>0$, let $\sigma^A : Q \to \mathbb{R}$ be defined by

$$\sigma^A(u,v) := \bar{\sigma}(u,v) + \mathcal{D}^A(u,v) \qquad \forall (u,v) \in Q.$$

By the above observation, if $(u,v),(u,v') \in \mathcal{Q}$ (resp. $(u,v),(u',v) \in \mathcal{Q}$) with $v' \neq v$ (resp. $u' \neq u$) satisfy $\nabla_u\sigma^A(u,v) = \nabla_u\sigma^A(u,v')$ (resp. $\nabla_v\sigma^A(u,v) = \nabla_v\sigma^A(u',v)$) and $A > L\sqrt{2n}$ where L is the maximum of $|\nabla_u\bar{\sigma}|, |\nabla_u\bar{\sigma}|$ over Q, then $\nabla_u\bar{\sigma}(u,v) = \nabla_u\bar{\sigma}(u,v')$ (resp. $\nabla_v\bar{\sigma}(u,v) = \nabla_v\bar{\sigma}(u',v)$). In conclusion, for $A>0$ large enough the $(\mathcal{Q},\sigma^A)$-chains are $(\mathcal{S},\bar{\sigma})$-chains. We conclude easily the lemma in the case $\lambda=1$. The general result for$\lambda>0$ follows by setting $\sigma^\lambda(u,v) := \sigma(u/\lambda, v/\lambda)$. □

To conclude, repeating the proof of Lemma 4.4 with l^∞-balls, the set B admits a countable tiling by l^∞-balls (balls for the norm $\|\cdot\|_\infty$ in $\mathbb{R}^n \times \mathbb{R}^n$). In other words, there are a family $\{z_j\}_{j\in\mathbb{N}}$ in $\mathrm{Int}(B)$ and a family of positive real numbers $\{r_j\}_{j\in\mathbb{N}}$ such that

$$\mathrm{Int}(B) = \bigcup_{j\in\mathbb{N}} B^\infty(z_j,r_j)$$

and

$$\mathrm{Int}\left(B^\infty(z_j,r_j)\right) \cap \mathrm{Int}\left(B^\infty(z_{j'},r_{j'})\right) = \emptyset \qquad \forall j \neq j' \in \mathbb{N}.$$

Each ball $B^\infty(z_j,r_j)$ can be translated to λQ with $\lambda_j = 2r_j$. So we can define on each $B^\infty(z_j,r_j)$ a cost k_j of the form $\nu_j\sigma^{\lambda_j}(\cdot - z_j)$ with $\nu_j \in (0,\lambda_j)$ and associated with some null set $\mathcal{N}^j \subset B^\infty(z_j,r_j)$ containing $\partial B^\infty(z_j,r_j)$ in such a way that any chain for the cost $h+k$ with

$$k(u,v) := k_j(u,v) \qquad \forall (u,v) \in B^\infty(z_j,r_j),\, \forall j \in \mathbb{N},$$

which is contained in the set $\mathrm{Int}(B) \setminus (\cup_j \mathcal{N}^j)$ is indeed contained in some $B^\infty(z_j, r_j)$ and so is a chain for the cost k_j in $B^\infty(z_j, r_j) \setminus \mathcal{N}^j$. By construction (P1) and (P3) are satisfied. Moreover, by taking the ν_j's sufficiently small, we get the property (P3). This concludes the proof of Theorem 1.5.

5 Proofs of the results stated in Section 2

In this section we first prove Theorems 2.6 and 2.10. Then we state and prove a generalization of Theorem 2.12 for which is applicable to unbounded cost functions.

5.1 Proof of Theorem 2.6

The result will follow from the existence of a universally measurable numbered limb system together with an extension of Theorem 2.4 (Theorem C.2) given in Appendix C. Let $\{\mathcal{S}_i\}_{i\in\mathbb{N}}$ be a partition of $\mathcal{S}$ into countably many Borel sets and and a sequence $\{L_i\}_{i\in\mathbb{N}}$ of integers ≥ 2 such that

$$\mathcal{O}^{c,\mathcal{S}}_{\geq L_i}(x,y) \cap \mathcal{S}_i = \emptyset \qquad \forall (x,y) \in \mathcal{S}_i,\, \forall i \in \mathbb{N}. \tag{5.1}$$

Let us first prove the result in the case where $L_i \equiv 2$. Then (5.1) becomes

$$\mathcal{O}^{c,\mathcal{S}}_{\geq 2}(x,y) \cap \mathcal{S}_i = \emptyset \qquad \forall (x,y) \in \mathcal{S}_i,\, \forall i \in \mathbb{N}. \tag{5.2}$$

Recall that the maps $F^1, F^2 : \mathcal{P}(M \times N) \to \mathcal{P}(M \times N)$ are defined by

$$F^1(A) := \left(\pi^1\right)^{-1}\left[\pi^1(A)\right] \quad \text{and} \quad F^2(A) = \left(\pi^2\right)^{-1}\left[\pi^2(A)\right] \qquad \forall A \subset M \times N.$$

For every $i \in \mathbb{N}$, we define a sequence of sets $\{\mathcal{S}_i^j\}_{j\in\mathbb{N}}$ recursively by $\mathcal{S}_i^0 := \mathcal{S}_i$ and for every $j \in \mathbb{N}$,

$$\begin{aligned} \mathcal{S}_i^{j+1} &:= F^2(\mathcal{S}_i^j) \cap \mathcal{S} \quad \text{if} \quad j \text{ is even},\\ \mathcal{S}_i^{j+1} &:= F^1(\mathcal{S}_i^j) \cap \mathcal{S} \quad \text{if} \quad j \text{ is odd}. \end{aligned}$$

By construction, all the sets $\mathcal{S}_i^j$ with $i, j \in \mathbb{N}$ are analytic so universally measurable (see [9, Lemma 15]). If for every $i \in \mathbb{N}$ and every $l \in \mathbb{N}^*$ we denote respectively by H_i^l and V_i^l the set of pairs that can be reached by a $\mathcal{S}$-chain of length l starting from $\mathcal{S}_i$ horizontally (resp. vertically), then we have

$$\mathcal{S}_i^{j+1} = \mathcal{S}_i^j \cup \left(V_i^{j+1} \cup H_i^{j+2}\right) = \bigcup_{l=1}^{j+1} \left(V_i^l \cup H_i^{l+1}\right) \qquad \forall i, j \in \mathbb{N}.$$

Moreover, thanks to (5.2), it can be shown (see [9, Lemma 7]) that for every $i \in \mathbb{N}$ and for every $l \in \mathbb{N}^*$, if l is even the set $V_i^l \cup H_i^{l+1}$ is the antigraph of a function

$f_i^l : \mathrm{Dom}(f_i^l) \subset N \to \mathrm{Ran}(f_i^l) \subset M$, if l is odd it is the graph of a function $f_i^l : \mathrm{Dom}(f_i^l) \subset M \to \mathrm{Ran}(f_i^l) \subset N$, and we have

$$\mathrm{Ran}\left(f_i^{2j+3}\right) \subset \mathrm{Dom}\left(f_i^{2j+2}\right), \quad \mathrm{Ran}\left(f_i^{2j+2}\right) \subset \mathrm{Dom}\left(f_i^{2j+1}\right) \qquad \forall i \in \mathbb{N}, \forall j \in \mathbb{N}$$

and for every $i, j, j' \in \mathbb{N}$

$$j \neq j' \quad \Longrightarrow \quad \begin{cases} \mathrm{Dom}(f_i^{2j+1}) \cap \mathrm{Dom}(f_i^{2j'+1}) & = & \emptyset \\ \mathrm{Dom}(f_i^{2j+2}) \cap \mathrm{Dom}(f_i^{2j'+2}) & = & \emptyset. \end{cases}$$

Therefore, for every $i \in \mathbb{N}$, setting for every $j \in \mathbb{N}$,

$$I_i^{2j+1} := \mathrm{Dom}(f_i^{2j+2}) \quad \text{and} \quad I_i^{2j+2} := \mathrm{Dom}(f_i^{2j+2}) \subset N,$$

we get a decomposition of the set

$$\tilde{\mathcal{S}}_i := \bigcup_{j \in \mathbb{N}} \mathcal{S}_i^j,$$

as a numbered limb system whose limbs are the universally measurable sets

$$\mathcal{L}_i^{2j+1} := \mathrm{Graph}(f_i^{2j+1}) \quad \text{and} \quad \mathcal{L}_i^{2j+2} := \mathrm{Antigraph}(f_i^{2j+2}) \qquad \forall j \in \mathbb{N}.$$

The above construction works for each $i \in \mathbb{N}$ because, by (5.2), each $\mathcal{S}_i$ is a $\tilde{\mathcal{S}}_i$-rooting set. If two sets of the form $\tilde{\mathcal{S}}_i$ and $\tilde{\mathcal{S}}_{i'}$ were disjoint then the union of their numbered limb systems would give a numbered limb system for the union $\tilde{\mathcal{S}}_i \cup \tilde{\mathcal{S}}_{i'}$. We cannot apply this simple technique in our case, we proceed as follows. For every $i \in \mathbb{N}$, we set

$$\tilde{\mathcal{S}}_{\leq i} := \bigcup_{l=0}^{i} \tilde{\mathcal{S}}_l,$$

then we define a family of limbs $\{\mathcal{L}_j\}_{j \in \mathbb{N}^*}$ by

$$\mathcal{L}_j := \mathcal{L}_0^j \cup \left(\bigcup_{i \in \mathbb{N}^*} \mathcal{L}_i^j \setminus \tilde{\mathcal{S}}_{\leq i-1} \right) \qquad \forall j \in \mathbb{N}^*.$$

By construction, all the $\mathcal{L}_j$'s are universally measurable sets, the $\mathcal{L}_j$'s with j odd are graphs, because such $\mathcal{L}_j$ is the union of graphs with disjoint domains (if two points $(x,y) \in \mathcal{L}_i^j$ and $(x',y') \in \mathcal{L}_{i'}^j \setminus \tilde{\mathcal{S}}_{\leq i'-1}$ with $i < i'$ have the same projection with respect to π^1, that is $x = x'$, then $x', y')$ belongs to the $\mathcal{S}$-orbit of (x,y) so it belongs to $\tilde{\mathcal{S}}_{\leq i'-1}$), and in the same way the $\mathcal{L}_j$'s with j even are antigraphs. Since the domains of the graphs and antigraphs appearing in the $\mathcal{L}_j$'s we check easily that the family $\{\mathcal{L}_j\}_{j \in \mathbb{N}^*}$

provides a decomposition of $\mathcal{S}$ into a numbered limb system with universally measurable limbs.

Let us now treat the general case corresponding to the assumption (5.1). We just need to show how to construct a numbered limb systems with universally measurable limbs for the sets $\tilde{\mathcal{S}}_i$ defined above. As a matter of fact, as soon as we have those decomposition, we can repeat the above proof to get a numbered limb system for $\mathcal{S}$. We show how to proceed for $i = 0$, without loss of generality we may assume that $L_0 = 2\ell$ for some $\ell \in \mathbb{N}^*$. The sequence of sets $\{\mathcal{S}_0^j\}_{j\in\mathbb{N}}$ defined recursively by $\mathcal{S}_0^0 := \mathcal{S}_0$ and for every $j \in \mathbb{N}$,

$$\mathcal{S}_0^{j+1} := F^2(\mathcal{S}_0^j) \cap \mathcal{S} \quad \text{if} \quad j \text{ is even,}$$

$$\mathcal{S}_0^{j+1} := F^1(\mathcal{S}_0^j) \cap \mathcal{S} \quad \text{if} \quad j \text{ is odd,}$$

verifies

$$\mathcal{S}_0^{j+1} = \mathcal{S}_0^j \cup \left(V_0^{j+1} \cup H_0^{j+2}\right) = \bigcup_{l=1}^{j+1} \left(V_0^l \cup H_0^{l+1}\right) \qquad \forall i, j \in \mathbb{N},$$

where for every $l \in \mathbb{N}^*$ we denote respectively by H_i^l and V_i^l the set of pairs that can be reached by a $\mathcal{S}$-chain of minimal length l starting from $\mathcal{S}_i$ horizontally (resp. vertically). The definitions of $V_0^l \cup H_0^{l+1}$ call for comment. Since $\mathcal{S}_0$ is not necessary a rooting set for $\tilde{S}_0 := \cup_{j\in\mathbb{N}}\mathcal{S}_0^j$, a point (x, y) in V_0^l for some $l > 1$ may admit a $\mathcal{S}$-chain $((x_1, y_1), \ldots, (x_{l'}, y_{l'}))$ with $(x_1, y_1) = (x, y)$, $(x_{l'}, y_{l'}) \in \mathcal{S}_0$ and $l' > 1$ but in this case we have necessarily $(x_l, y_l) \in \mathcal{S}_0$ and $(x_1, y_1), \ldots, (x_{l-1}, y_{l-1}) \notin \mathcal{S}_0$. By (5.1), for every $l \geq L_0/2 = \ell$ odd the set $V_0^l \cup H_0^{l+1}$ is a graph and for every $l \geq \ell$ even this set is an antigraph. So we will get a partial numbered limb systems. Then to recover the plan over $I_0 \times I_1$ we will proceed as in the proof of Theorem 1.4 in [14]. For every $(x, y) \in \mathcal{S}$, we define $\ell(x, y)$ as the supremum of all natural numbers $L \in \mathbb{N}^*$ such that there is at least one $\mathcal{S}$-chain $((x_1, y_1), \ldots, (x_L, y_L))$ with $(x_L, y_L) = (x, y)$. Moroever we say that $((x_1, y_1), \ldots, (x_L, y_L))$ has an horizontal end if $y_L = y_{L-1}$ and a vertical end if $x_L = x_{L-1}$. Then we set

$$\mathcal{S}_L := \left\{(x, y) \in \mathcal{S} \,|\, \ell(x, y) \geq L\right\}$$

and denote by $\mathcal{S}_L^h$ (resp. $\mathcal{S}_L^v$) the set of pairs $(x, y) \in \mathcal{S}_L$ such that there is a $\mathcal{S}$-chain $((x_1, y_1), \ldots, (x_L, y_L))$ of length L with horizontal end at $(x, y) = (x_L, y_L)$ (resp. with vertical end at $(x, y) = (x_L, y_L)$). Let d_M, d_N be Riemannnian distances on M and N, for every $p \geq 1$ and every integer $L \geq 2$, denote by S_p^L the closed set of L-tuples $((x_1, y_1), \ldots, (x_L, y_L))$ in $(M \times N)^L$ such that for every $l = 1, \ldots, L-1$ either

$$x_l = x_{l+1}, \quad y_l \neq y_{l+1} = y_{\min\{L, l+2\}} \quad \text{and} \quad d_N(y_l, y_{l+1}) \geq 1/p,$$

or

$$y_l = y_{l+1}, \quad x_l \neq x_{l+1} = x_{\min\{L,l+2\}} \quad \text{and} \quad d_M(x_l, x_{l+1}) \geq 1/p.$$

A pair (x, y) belongs to $\mathcal{S}_L$ if and only if there is $p \geq 1$ such that it is the image of a L-tuple in the Borel set $S_p^L \cap (\mathcal{S})^L$ by the projection

$$((x_1, y_1), \ldots, (x_L, y_L)) \in (M \times N)^L \longmapsto (x_L, y_L) \in M \times N.$$

Thus $\mathcal{S}_L$ is an analytic set in $M \times N$. Specifying the end at L, we show in the same way that $\mathcal{S}_L^h, \mathcal{S}_L^v$ are analytic. Then proceeding as in the proof of Theorem 1.4 in [14], we can show that the sets defined by

$$\begin{aligned} G_1 &:= E_1 \cup E_2^{h-}, \\ G_{2i} &:= E_{2i}^v \cup E_{2i+1}^{v-} \quad \forall i \in \mathbb{N}^*, \\ G_{2i+1} &:= E_{2i+2}^h \cup E_{2i+1}^{h-} \quad \forall i \in \mathbb{N}^*, \end{aligned}$$

where $E_1 := \mathcal{S}_1 \setminus \mathcal{S}_2$ and for $k \geq 2$ $E_k^h := (\mathcal{S}_k \setminus \mathcal{S}_{k+1}) \cap \mathcal{S}_k^h$, $E_k^v := (\mathcal{S}_k \setminus \mathcal{S}_{k+1}) \cap \mathcal{S}_k^v$, $E_k^{h-} := (\mathcal{S}_k \setminus \mathcal{S}_{k+1}) \setminus \mathcal{S}_k^v$ and $E_k^{v-} := (\mathcal{S}_k \setminus \mathcal{S}_{k+1}) \setminus \mathcal{S}_k^h$, form the graphs and the antigraphs of a numbered limb system with universally measurable limbs. We conclude by Theorem C.2.

5.2 Proof of Theorem 2.10

The main idea of the proof is to show that a (c, μ, ν)-minimizing set $\mathcal{S}$ which is c-extreme is indeed a Borel strongly disjoint union of a graph and an antigraph. Then we conclude the desired results from Theorem 2.8.
We apply Lemma B.7 and get Borel measurable functions $h : \mathrm{Dom}(h) \subset \bar{M} \to N$ and $g : \mathrm{Dom}(g) \subset \bar{N} \to M$ satisfying the following properties:

(i) There exists a full measure subset $\tilde{M}$ of $\bar{M}$ such that

$$\mathcal{S}' := \mathcal{S} \cap \left(\tilde{M} \times \bar{N}\right) \subset \mathrm{Graph}(h) \cup \mathrm{Antigraph}(g).$$

(ii) $\mathrm{Graph}(h) \cap \mathrm{Antigraph}(g) = \emptyset$.

(iii) $\mathrm{Dom}(h) \subset D(\mathcal{S}, c)$ and $h = f_{\mathcal{S},c}$ on $\mathrm{Dom}(h)$ where

$$D(\mathcal{S}, c) := \Big\{x \in M \mid \{\text{argmax } c(x, y);\, y \in \Gamma_{\mathcal{S}}(x)\} \text{ is a singleton}\Big\}.$$

In order to apply Theorem 2.8, we need a Borel measurable bounded function $\theta : N \to \mathbb{R}$ such that $\theta(h \circ g(y)) > \theta(y)$ for every $y \in \mathrm{Dom}(h \circ g)$. To this aim we consider the continuous potentials $\psi : M \to \mathbb{R}$ and $\phi : M \to \mathbb{R}$ given by duality and satisfying

$$\psi(x) = \max\Big\{\phi(y) - c(x,y) \,|\, y \in N\Big\} \qquad \forall x \in M,$$

$$\phi(y) = \min\Big\{\psi(x) + c(x,y) \,|\, x \in M\Big\} \qquad \forall y \in N,$$

$$\text{and} \quad \mathcal{S} \subset \partial_c\psi := \Big\{(x,y) \in M \times N \,|\, c(x,y) = \phi(y) - \psi(x)\Big\}.$$

We shall prove that $\phi\big(h \circ g(y)\big) > \phi(y)$ for every $y \in \mathrm{Dom}(h \circ g)$. So let us consider $y \in \mathrm{Dom}(h \circ g)$ and set $x := g(y) \in Dom(h)$. Note that $y \neq h(x)$, as otherwise

$$\big((x, h(x)\big) = \big(g(y), y\big) \in \mathrm{Graph}(h) \cap \mathrm{Antigraph}(g),$$

which contradicts (ii) in Lemma B.7. Since $(x,y), \big(x, h(x)\big) \in \mathcal{S}$, we also have that

$$c(x,y) = \phi(y) - \psi(x) \quad \text{and} \quad c\big(x, h(x)\big) = \phi\big(h(y)\big) - \psi(x). \tag{5.3}$$

On the other hand $x \in \mathrm{Dom}(h) \subset D(\mathcal{S}, c)$ and $h(x) = f_{\mathcal{S},c}(x) = \mathrm{argmax}_{y \in \Gamma_{\mathcal{S}}(x)} c(x,y)$ from which we obtain

$$c\big(x, h(x)\big) > c(x,y).$$

This together with (5.3) yield that

$$\phi\big(h \circ g(y)\big) = \phi\big(h(x)\big) = c\big(x, h(x)\big) + \psi(x) > c(x,y) + \psi(x) = \phi(y).$$

Now by setting $\theta = \phi$, all assumptions of Theorem 2.8 are satisfied and we can conclude easily.

5.3 Proof of Theorem 2.12

Here we shall prove a generalized version of Theorem 2.12 for which the cost function is allowed to be even unbounded.

Let M and N be Polish spaces. Fix a Borel measurable cost $c : M \times N \to (-\infty, +\infty]$, and two Borel probability measures μ, ν respectively on M and N.

Definition 5.1 (Strongly (c, μ, ν)-minimizing sets). *Say that a Borel measurable set $\mathcal{S} \subset M \times N$ is a strongly (c, μ, ν)-minimizing set provided the following conditions hold;*

1. *There exist Borel measurable functions* $\phi : N \to [-\infty, +\infty)$ *and* $\psi : M \to (-\infty, +\infty]$ *such that*

$$\psi(x) = \sup_{y \in N}\{\phi(y) - c(x,y)\} \qquad \& \qquad \phi(y) = \inf_{x \in M}\{c(x,y) + \psi(x)\},$$

and

$$\mathcal{S} \subseteq \big\{(x,y) \in M \times N;\, \phi(y) - \psi(x) = c(x,y)\big\}.$$

2. *There exists* $\gamma \in \Pi(\mu, \nu)$ *with* $\gamma(S) = 1$. *Also, there exists a* μ *full measure Borel set* $M_1 \subset M$ *and a* ν *full measure Borel set* $N_1 \subset N$ *that* $M_1 \subseteq \pi^M(\mathcal{S})$ *and* $N_1 \subseteq \pi^N(\mathcal{S})$.

As before, for a strongly (c, μ, ν)-minimizing set $\mathcal{S}$, we define the set-valued maps $\Gamma_{\mathcal{S}} : M \to 2^N$ by

$$\Gamma_{\mathcal{S}}(x) = \Big\{y \in N \mid (x,y) \in \mathcal{S}\Big\},$$

and $f_{\mathcal{S},c} : M \to 2^N$ by

$$f_{\mathcal{S},c}(x) := \text{argmax}\Big\{c(x,y) \,|\, y \in \Gamma_{\mathcal{S}}(x)\Big\} = \Big\{y \in \Gamma_{\mathcal{S}}(x) \,|\, c(x,y) = \max_{z \in \Gamma_{\mathcal{S}}(x)} c(x,z)\Big\}.$$

Definition 5.2 (Strongly c-extreme sets). *We say that* $\mathcal{S} \subset M \times N$ *is strongly c-extreme if there are full measure Borel sets* $\bar{M}$ *of* M *and* $\bar{N}$ *of* N *such that the following assertions hold:*

1. *for each* $x \in Dom\big(\Gamma_{\mathcal{S}}\big) \cap \bar{M}$ *the set* $\{\text{argmax}\, c(x,y);\, y \in \Gamma_{\mathcal{S}}(x)\}$ *is non-empty;*
2. *for all distinct points* $x_1, x_2 \in \bar{M}$,

$$\Big\{\Gamma_{\mathcal{S}}(x_1) \setminus \{y_1\}\Big\} \cap \Big\{\Gamma_{\mathcal{S}}(x_2) \setminus \{y_2\}\Big\} \cap \bar{N} \;\; = \;\; \emptyset,$$

for all $y_1 \in f_{\mathcal{S},c}(x_1)$ *and* $y_2 \in f_{\mathcal{S},c}(x_2)$.

Note that if c is continuous and M, N are compact manifolds then a (c, μ, ν)-minimizing set (resp. c-extreme set) is also strongly (c, μ, ν)-minimizing (resp. strongly c-extreme).

We have the following result which is a generalization of Theorem 2.12.

Theorem 5.3. *Assume that* M, N *are Polish spaces equipped with Borel probability measures* μ *and* ν *respectively. Let* $c : M \times N \to [0, \infty]$ *be Borel measurable, lower semi-continuous and* $\mu \otimes \nu -$*a.e. finite. Assume that* $\mathcal{S} \subset M \times N$ *is a strongly* (c, μ, ν)*-minimizing set. If there exists a Borel ordered partition* P *of* N *such that* $\mathcal{S}$ *is* (c, P)*-extreme, then there is a unique* $\gamma \in \Pi(\mu, \nu)$ *such that* $\gamma(\mathcal{S}) = 1$.

In order to mimic the proof of Theorem 5.3, we need to construct a sequence of costs associated with the ordered partition $P = \{Y_i\}_i$ of N. Indeed, for each partition P, we introduce a new cost function c_P in such a way that c_P-extreme measures are $(c, P)-$extreme measures and vice-versa. Let $P = \{Y_i\}_{i=1}^{L}$ be a measurable ordered partition of Y where $L \in \mathbb{N} \cup \{+\infty\}$. Define the bounded function $c_P : M \times N \to \mathbb{R}$ by

$$c_P(x, y) = \begin{cases} \frac{c(x,y)}{2^{2i-1}(1+|c(x,y)|)} + \frac{1}{2^{2i-2}}, & \text{if } c(x, y) \in \mathbb{R} \\ \frac{1}{2^{2i-1}} + \frac{1}{2^{2i-2}}, & \text{if } c(x, y) = +\infty. \end{cases}$$

for $(x, y) \in M \times Y_i$. Since $Y_i \cap Y_j = \emptyset$ for all $i \neq j$ we have that c_P is well-defined. The measurability of c_P follows from the measurability of c and the fact that each Y_i is measurable.

Lemma 5.4. *The following assertions hold:*

1. $D(\mathcal{S}, c, P) = D(\mathcal{S}, c_P)$.
2. $f_{\mathcal{S},c,P} = f_{\mathcal{S},c_P}$.
3. *$\mathcal{S}$ is $(c, P)-$extreme if and only if $\mathcal{S}$ is strongly c_P-extreme.*

Proof. For $i < L$, take arbitrary elements $y_i \in Y_i$ and $y_{i+1} \in Y_{i+1}$. It follows that

$$\begin{aligned} c_P(x, y_i) &\geq -\frac{1}{2^{2i-1}} + \frac{1}{2^{2i-2}} \\ &> \frac{1}{2^{2i+1}} + \frac{1}{2^{2i}} \\ &\geq c_P(x, y_{i+1}). \end{aligned}$$

Thus, $c_P(x, y_i) > c_P(x, y_{i+1})$. This also shows that

$$c_P(x, y_i) > c_P(x, y_j) \qquad \forall i < j, \quad \forall y_i \in Y_i, \quad \forall y_j \in Y_j. \tag{5.4}$$

We shall now show that for each $x \in Dom(\Gamma_{\mathcal{S}})$,

$$\big\{ \operatorname{argmax} c_p(x, y);\, y \in \Gamma_{\mathcal{S}}(x) \big\} = \big\{ \operatorname{argmax} c(x, y);\, y \in \Gamma_{\mathcal{S}}(x) \cap Y_{i(x)} \big\}, \tag{5.5}$$

where $i(x) = \min\{i \in \mathbb{N};\, \Gamma_{\mathcal{S}}(x) \cap Y_i \neq \emptyset\}$ as in Definition 2.11. Take

$$y_0 \in \big\{ \operatorname{argmax} c_p(x, y);\, y \in \Gamma_{\mathcal{S}}(x) \big\}.$$

It follows from (5.4) that $y_0 \in Y_{i(x)}$. If now $c(x, y_0) < \sup_{y \in \Gamma_{\mathcal{S}}(x) \cap Y_{i(x)}} c(x, y)$ then there exists $y_1 \in \Gamma_{\mathcal{S}}(x) \cap Y_{i(x)}$ such that $c(x, y_0) < c(x, y_1)$. If $c(x, y_1) \neq +\infty$ then

$$c_P(x, y_1) = \frac{c(x, y_1)}{2^{2i(x)-1}(1 + |c(x, y_1)|)} + \frac{1}{2^{2i(x)-2}} > \frac{c(x, y_0)}{2^{2i(x)-1}(1 + |c(x, y_0)|)} + \frac{1}{2^{2i(x)-2}} = c_P(x, y_0),$$

and if $c(x, y_1) = +\infty$ then

$$c_P(x, y_1) = \frac{1}{2^{2i(x)-1}} + \frac{1}{2^{2i(x)-2}} > \frac{c(x, y_0)}{2^{2i(x)-1}(1 + |c(x, y_0)|} + \frac{1}{2^{2i(x)-2}} = c_P(x, y_0).$$

Thus, in both cases we have $c_P(x, y_1) > c_P(x, y_0)$ which contradicts the fact that

$$c_P(x, y_0) = \max_{y \in \Gamma_{\mathcal{S}}(x)} c_P(x, y).$$

Therefore, we must have

$$c(x, y_0) = \max_{y \in \Gamma_{\mathcal{S}}(x) \cap Y_{i(x)}} c(x, y),$$

from which we obtain

$$y_0 \in \left\{ \operatorname{argmax} c(x, y);\ y \in \Gamma_{\mathcal{S}}(x) \cap Y_{i(x)} \right\}.$$

We shall now prove the other direction, i.e.

$$\left\{ \operatorname{argmax} c(x, y);\ y \in \Gamma_{\mathcal{S}}(x) \cap Y_{i(x)} \right\} \subseteq \left\{ \operatorname{argmax} c_p(x, y);\ y \in \Gamma_{\mathcal{S}}(x) \right\}.$$

Take $y_0 \in \left\{ \operatorname{argmax} c(x, y);\ y \in \Gamma_{\mathcal{S}}(x) \cap Y_{i(x)} \right\}$. It follows that

$$c(x, y_0) = \max_{y \in \Gamma_{\mathcal{S}}(x) \cap Y_{i(x)}} c(x, y). \tag{5.6}$$

If $c_p(x, y_0) < \sup_{y \in \Gamma_{\mathcal{S}}(x)} c_p(x, y)$ then there exists $y_1 \in \Gamma_{\mathcal{S}}(x)$ such that $c_P(x, y_0) < c_P(x, y_1)$. It is easily seen that $y_1 \in Y_{i(x)}$ as otherwise $c_P(x, y_0) > c_P(x, y_1)$ due to (5.4). It now follows from (5.6) that $c(x, y_1) \leq c(x, y_0)$ and this inequality is indeed strict as $c_P(x, y_0) < c_P(x, y_1)$. We must then have $c(x, y_0), c(x, y_1)$ are real-valued. Thus,

$$\begin{aligned} \frac{c(x, y_1)}{2^{2i(x)-1}(1 + |c(x, y_1)|)} + \frac{1}{2^{2i(x)-2}} = c_P(x, y_1) &> c_P(x, y_0) \\ &= \frac{c(x, y_0)}{2^{2i(x)-1}(1 + |c(x, y_0)|)} + \frac{1}{2^{2i(x)-2}}. \end{aligned}$$

from which we obtain $c(x, y_1) > c(x, y_0)$ which leads to a contradiction.
Parts 1) and 2) then follow from (5.5). Part 3) is a direct consequence of parts 1) and 2). $\square$

Proof of Theorem 5.3. By Lemma 5.4, the set $\mathcal{S}$ is strongly c_P-extreme. It then follows from Lemma B.7 that there exist Borel measurable functions $h : Dom(h) \subseteq M \to N$ and $g : Dom(g) \subseteq N \to M$ satisfying the following properties:

(i) There exists a full measure subset $\tilde{M}$ of $\bar{M}$ such that

$$\mathcal{S}' := \mathcal{S} \cap \Big(\tilde{M} \times \bar{N}\Big) \subset \mathrm{Graph}(h) \cup \mathrm{Antigraph}(g).$$

(ii) $\mathrm{Graph}(h) \cap \mathrm{Antigraph}(g) = \emptyset$.

(iii) $\mathrm{Dom}(h) \subset D(\mathcal{S}, c_P)$ and $h = f_{\mathcal{S},c_P}$ on $\mathrm{Dom}(h)$ where

$$D(\mathcal{S}, c_P) := \Big\{x \in M \mid \{\text{argmax } c_P(x,y);\, y \in \Gamma_{\mathcal{S}}(x)\} \text{ is a singleton } \Big\}.$$

Since $\mathcal{S}$ is a strongly (c,μ,ν)-minimizing set, there exist Borel measurable functions $\phi : N \to [-\infty, +\infty)$ and $\psi : M \to (-\infty, +\infty]$ with $\phi(y) - \psi(x) \le c(x,y)$ on $M \times N$ such that

$$\mathcal{S} \subseteq \big\{(x,y) \in M \times N;\, \phi(y) - \psi(x) = c(x,y)\big\}.$$

Define $\theta : N \to \mathbb{R}$ by

$$\theta(y) = \begin{cases} \frac{\phi(y)}{2^{2i-1}(1+|\phi(y)|)} + \frac{1}{2^{2i-2}}, & \text{if } \phi(y) \in \mathbb{R} \\ \frac{1}{2^{2i-1}} + \frac{1}{2^{2i-2}}, & \text{if } \phi(y) = -\infty. \end{cases}$$

for $y \in Y_i$. Note that, by the same argument as in Lemma 5.4, if $i < j$ then for each $y_i \in Y_i$ and each $y_j \in Y_j$ we have that $\theta(y_i) > \theta(y_j)$. We shall now prove that

$$\theta\big(h \circ g(y)\big) > \theta(y) \qquad \forall y \in Dom(h \circ g).$$

Take $y_0 \in Dom(h \circ g)$ and assume that $g(y_0) = x_0$. It implies that $x_0 \in Dom(h) \subset D(\gamma, c_P)$ and $y_0 \in \Gamma_{\mathcal{S}}(x_0) \setminus \{h(x_0)\}$. We also have that $(x_0, y_0), \big(x_0, h(x_0)\big) \in \mathcal{S}$. It implies that there exist i, j such that $h(x_0) \in Y_i$ and $y_0 \in Y_j$. Since

$$h = f_{\mathcal{S},c_P} = f_{\mathcal{S},c,P}, \quad \text{on } Dom(h),$$

we must have $i \le j$. We now consider two cases $i < j$ and $i = j$.
Case I. If $i < j$ then it follows from the definition of the function θ that

$$\theta\big(h \circ g(y)\big) = \theta\big(h(x_0)\big) > \theta(y_0), \qquad (\text{since } y_0 \in \mathcal{S}_j \text{ and } h(x_0) \in \mathcal{S}_i),$$

as desired.
Case II. If $i = j$, we have that

$$\phi(y_0) - \psi(x_0) = c(x_0, y_0) \qquad \& \qquad \phi\big(h(x_0)\big) - \psi(x_0) = c\big(x_0, h(x_0)\big). \tag{5.7}$$

On the other hand

$$h(x_0) = \underset{y \in \Gamma_{\mathcal{S}}(x_0) \cap Y_{i(x)}}{\text{argmax}} c(x_0, y),$$

from which we obtain $c\big(x_0, h(x_0)\big) > c(x_0, y_0)$. This together with (5.7) yield that

$$\phi\big(h \circ g(y_0)\big) = \phi\big(h(x_0)\big) = c\big(x_0, h(x_0)\big) + \psi(x_0) > c(x_0, y_0) + \psi(x_0) = \phi(y_0).$$

This in fact implies that $\theta\big(h \circ g(y_0)\big) > \theta(y_0)$. The result now follows from Theorem 2.8. $\square$

6 Uniqueness of optimal plans for infinite chains and non-compact settings

In this section we shall provide some interesting applications of Theorem 5.3. Let M and N be complete separable Borel metric spaces and let $c : M \times N \to [0, +\infty]$ be a lower semi-continuous Borel measurable function. For Borel probability measures μ on M and ν on N, consider the problem

$$\inf\Big\{\int_{M\times N} c(x,y)\, d\pi; \pi \in \Pi(\mu,\nu)\Big\}. \qquad (MK)$$

The following result is rather standard.

Lemma 6.1. *Assume that M, N are Polish spaces equipped with Borel probability measures μ, ν, that $c : M \times N \to [0, \infty]$ is lower semi-continuous, Borel measurable and $\mu \otimes \nu-$a.e. finite and, there exists a finite transport plan. Then there exist Borel measurable functions $\phi : N \to [-\infty, +\infty)$ and $\psi : M \to (-\infty, +\infty]$ with*

$$\psi(x) = \sup_{y\in N}\{\phi(y) - c(x,y)\} \qquad \& \qquad \phi(y) = \inf_{x\in M}\{c(x,y) + \psi(x)\}, \tag{6.1}$$

such that if γ is an optimal plan of (MK) then γ is concentrated on

$$\big\{(x,y) \in M \times N;\ \phi(y) - \psi(x) = c(x,y)\big\}.$$

We refer to Corollary 1.2 in [2] for the proof. As a direct consequence of the above Lemma we have the following result.

Corollary 6.2. *Under the assumptions of Lemma 6.1, there exists a closed strongly (c, μ, ν)-minimizing set $\mathcal{S} \subset M \times N$ such that $\gamma \in \Pi(\mu, \nu)$ is a solution of (MK) if and only if $\gamma(S) = 1$.*

6.1 Infinite chains and uniqueness

Let M and N be smooth closed manifolds of dimension $n > 1$ and $c : M \times N \to \mathbb{R}$ a Lipschitz cost function. We have the following definitions.

Definition 6.3 (**Right-Maximal chains**). *Let $\mathcal{S} \subset M \times N$ and $(x_1, y_1) \in \mathcal{S}$. Say that a finite $(c, \mathcal{S})$-chain $\{(x_1, y_1), ...(x_l, y_l)\}$ starting from (x_1, y_1) is Right-Maximal if for each $(\bar{x}, \bar{y}) \in \mathcal{S}$ the following set is not a $(c, \mathcal{S})$-chain,*

$$\{(x_1, y_1), ..., (x_l, y_l), (\bar{x}, \bar{y})\}.$$

Definition 6.4 (**Maximal infinite chains**). *Let $\mathcal{S} \subset M \times N$. An infinite $(c, \mathcal{S})$-chain*

$$T = \{(x_1, y_1), (x_2, y_2), (x_3, y_3), ...\},$$

is said to be maximal if for each $(\bar{x}, \bar{y}) \in \mathcal{S}$ the following set is not a $(c, \mathcal{S})$-chain

$$\{(\bar{x}, \bar{y}), (x_1, y_1), (x_2, y_2), (x_3, y_3), ...\}.$$

Definition 6.5 (**The characteristic value**). *Let $\mathcal{S} \subset M \times N$. For each $y \in N$, we denote by $\ell_{\mathcal{S}}(y)$ the characteristic value of y with respect to $\mathcal{S}$ and we define it as follows:*
For each pair $(x, y) \in \mathcal{S}$, let $\mathcal{RM}(x, y)$ be the set of all finite Right-Maximal $(c, \mathcal{S})$-chains starting from (x, y). For each $y \in N$, let $\mathcal{R}(y) = \cup_{x \in M} \mathcal{RM}(x, y)$ and define the characteristic value of y by

$$\ell_{\mathcal{S}}(y) = \sup_{R \in \mathcal{R}(y)} length(R).$$

If $\mathcal{R}(y) = \emptyset$, we conventionally define $\ell_{\mathcal{S}}(y) = 1$.

Note that if $\ell_{\mathcal{S}}(y) = 1$, then one of the following situation occurs:
1) For each $x \in M$, there is no $(c, \mathcal{S})$-chain starting from (x, y).
2) If for some $x \in M$, there is a $(c, \mathcal{S})$-chain R starting from (x, y) then R is an infinite $(c, \mathcal{S})$-chain.

We would also like to emphasize that if $1 < \ell_{\mathcal{S}}(y) < \infty$ for some $y \in N$, then it does not necessarily mean that any $(c, \mathcal{S})$-chain starting from (x, y), for some $x \in M$, has a finite length. Indeed, we can have infinite $(c, \mathcal{S})$-chains starting from (x, y). To be more precise, let

$$T = \{(x_1, y_1), (x_2, y_2), ...\},$$

be a Maximal infinite $(c, \mathcal{S})$-chain such that $(x, y) \in T$ and $1 < \ell_{\mathcal{S}}(y) < \infty$. Assume that $(x_i, y_i) = (x, y)$ for some $i \in \mathbb{N}$. Then we must have that $i \leq \ell_{\mathcal{S}}(y)$.

Definition 6.6 (**Infinite cycles**). *Recall that a* $(c, \mathcal{S})$*-chain* $\{(x_1, y_1), ...(x_l, y_l)\}$ *of length* $l \geq 5$ *is called cyclic (or finitely cyclic) if* $(x_1, y_1) = (x_l, y_l)$. *We say that a* $(c, \mathcal{S})$*-chain is an infinitely cyclic chain if it consists of two infinite* $(c, \mathcal{S})$*-chains*

$$\{(x_1, y_1), (x_2, y_2), ...\} \qquad \& \qquad \{(\bar{x}_1, \bar{y}_1), (\bar{x}_2, \bar{y}_2), ...\},$$

such that the following set is also a $(c, \mathcal{S})$*-chain,*

$$\{(\bar{x}_2, \bar{y}_2), (\bar{x}_1, \bar{y}_1), (x_1, y_1), (x_2, y_2)\}.$$

Basically, an infinitely cyclic $(c, \mathcal{S})$-chain is formed by gluing two infinite $(c, \mathcal{S})$-chains so that the resulting chain is also a $(c, \mathcal{S})$-chain.

Definition 6.7 (**$(c, \mathcal{S})$-cycle**). *Any* $(c, \mathcal{S})$*-chain that is finitely or infinitely cyclic is said to be a* $(c, \mathcal{S})$*-cycle.*

Before stating our result we recall the following definition from [16].

Definition 6.8 (**generalized-twist condition:**). *Let* $c : M \times N \to \mathbb{R}$ *be a Lipschitz function. We say that c satisfies the generalized-twist condition if for each* $x_0 \in M$ *and* $y_0 \in N$ *the set*

$$\Big\{y \in N;\ \tilde{D}_x^- c(x_0, y) \cap \tilde{D}_x^- c(x_0, y_0) \neq \emptyset\Big\},$$

is finite.

We remark that if $dim(M) = dim(N)$ and the cost function c is of class C^2, then the non-degeneracy condition,

$$det\Big(\frac{\partial^2}{\partial y \partial x} c(x, y)\Big) \neq 0, \quad \forall (x, y) \in M \times N,$$

implies the generalized-twist condition above (See proposition 1.1 in [16] for the proof). Here we state our result regarding $(c, \mathcal{S})$-chains.

Theorem 6.9. *Let* M *and* N *be smooth closed manifolds of dimension* $n > 1$ *and* $c : M \times N \to \mathbb{R}$ *be a Lipschitz cost function satisfying the generalized twist condition. Let* $\mathcal{S}$ *be a* $(c, \mu, \nu)-$*minimizing set. If there is no* $(c, \mathcal{S})$*-cycles (finite or infinite) and, if* $\ell_{\mathcal{S}}(y) < \infty$ *for all* $y \in N$ *then for any pair* μ, ν *of probability measures respectively on* M *and* N *which are both absolutely continuous with respect to Lebesgue, there is a unique optimal plan between* μ *and* ν.

We shall need some preliminary results before proving Theorem 6.9. For each $L \in \mathbb{N}$, set

$$N_L = \big\{y \in N;\ \ell_{\mathcal{S}}(y) = L\big\}.$$

Lemma 6.10. *For each $L \geq 1$, the set N_L is a Borel measurable subset of N.*

Proof. One can use a similar argument as in ([14], Lemma 4.4 and Corollary 4.5) to prove this lemma. Here we just sketch the proof. Let $\mathcal{S}_L(x,y)$ be the set of all Right-Maximal $(c,\mathcal{S})$-chains starting from (x,y) with length at least L and define

$$\mathcal{S}_L = \{(x,y) : \mathcal{S}_L(x,y) \neq \emptyset\}.$$

Note that

$$N_L = \pi^N\Big(\mathcal{S}_L \setminus \mathcal{S}_{L+1}\Big).$$

Let $L \geq 3$ be an odd number. Denote by $\mathcal{S}_L^h$ (resp. $\mathcal{S}_L^v$) the set of pairs in $\mathcal{S}_L$ such that there exists a chain $\{(x_1,y_1),...(x_L,y_L)\}$ in $\mathcal{S}_L(x_1,y_1)$ with $y_1 = y_2$ (resp. $x_1 = x_2$). Note that $\mathcal{S}_L = \mathcal{S}_L^h \cup \mathcal{S}_L^v$. We show that $\mathcal{S}_L^v$ is measurable. The measurability of $\mathcal{S}_L^h$ follows by a similar argument.
Let $\tilde{\mathcal{S}}_L(x,y)$ be the set of all $(c,\mathcal{S})$-chains starting from (x,y) with length at least L and define

$$\tilde{\mathcal{S}}_L = \big\{(x,y) : \tilde{\mathcal{S}}_L(x,y) \neq \emptyset\big\}.$$

Denote by $\tilde{\mathcal{S}}_L^h$ (resp. $\tilde{\mathcal{S}}_L^v$) the set of pairs in $\tilde{\mathcal{S}}_L$ such that there exists a chain

$$\{(x_1,y_1),...(x_L,y_L)\},$$

in $\mathcal{S}_L(x_1,y_1)$ with $y_1 = y_2$ (resp. $x_1 = x_2$). Note that by Lemma 4.4 in [14] both $\tilde{\mathcal{S}}_L^v$ and $\tilde{\mathcal{S}}_L^h$ are Borel measurable. Define

$$O := \mathcal{S} \setminus \tilde{\mathcal{S}}_2^v.$$

Endow the manifolds M and N with Riemannian distances d_M and d_N. For every integer p, denote by S_p the set of L-tuples

$$T = \{(x_1,y_1),...,(x_L,y_L)\} \subset (M \times N)^{L-1} \times O$$

such that T is a $(c,\mathcal{S})$-chain and for every $l = 1,\ldots,L-1$ there holds,

$$\text{for } l \text{ odd: } x_l = x_{l+1} \quad \text{and} \quad d_N(y_l,y_{l+1}) \geq \frac{1}{p},$$

$$\text{for } l \text{ even: } y_l = y_{l+1} \quad \text{and} \quad d_M(x_l,x_{l+1}) \geq \frac{1}{p}.$$

Note that $(x,y) \in \mathcal{S}_L^v$ if and only if $(x,y) \in Proj_1(S_p)$ for some integer p. Here, $Proj_1 : (M \times N)^L \to M \times N$ is defined by

$$Proj_1\{(x_1,y_1),...,(x_L,y_L)\} = (x_1,y_1).$$

One can now use similar ideas as in ([14], Lemma 4.4 and Corollary 4.5) to complete the proof. $\square$

Proof of Theorem 6.9. By Lemma 6.10, each N_l is measurable. Since $\ell_{\mathcal{S}}(y) < \infty$ for every $y \in N$ we have that $P = \{N_l\}$ is an ordered partition of N. We shall show that $\mathcal{S}$ is $(c, P)-$extreme and then the result follows from Theorem 5.3. Since μ and ν are absolutely continuous there exist full measure subsets $\bar{M} \subseteq M$ and $\bar{N} \subseteq N$ such that the Kantorovich potentials ϕ and ψ given in Lemma 2.13 are differentiable on $\bar{M}$ and $\bar{N}$ respectively. Therefore, if $(x, y_1), (x, y_2) \in \mathcal{S}$ with $x \in \bar{M}$ we must have

$$-D\psi(x) \in \tilde{D}_x^- c(x, y_1) \cap \tilde{D}_x^- c(x, y_2),$$

since $\phi(y) - \psi(z) \leq c(z, y)$ for all $(z, y) \in M \times N$ and

$$\psi(y_i) - \phi(x) = c(x, y_i), \qquad i = 1, 2.$$

Similarly, if $(x_1, y), (x_2, y) \in \mathcal{S}$ with $y \in \bar{N}$ then we must have

$$D\phi(y) \in \tilde{D}_y^- c(x_1, y) \cap \tilde{D}_y^- c(x_2, y).$$

As in Definition 2.11, define the maps $i : Dom(f_{\mathcal{S},c}) \to \mathbb{N}$ and $f_{\mathcal{S},c,P} : M \to 2^N$ by

$$i(x) = \min\{l \in \mathbb{N}; \Gamma_{\mathcal{S}}(x) \cap N_l \neq \emptyset\}$$

and

$$f_{\mathcal{S},c,P}(x) = \text{argmax}\left\{c(x, y);\, y \in \Gamma_{\mathcal{S}}(x) \cap N_{i(x)}\right\}.$$

Note that since c satisfies the generalized twist condition, the set $\Gamma_{\mathcal{S}}(x)$ is finite for $\mu-$almost every $x \in X$ and therefore,

$$\text{argmax}\left\{c(x, y);\, y \in \Gamma_{\mathcal{S}}(x) \cap N_{i(x)}\right\},$$

is non-empty for $\mu-$almost every $x \in X$. To prove condition (ii) in Definition 2.11, take distinct points $x, \bar{x} \in \bar{M}$. Take $\bar{w} \in f_{\gamma,c,P}(\bar{x})$ and $w \in f_{\gamma,c,P}(x)$. We show that

$$\left\{\Gamma_{\mathcal{S}}(x) \setminus \{w\}\right\} \cap \left\{\Gamma_{\mathcal{S}}(\bar{x}) \setminus \{\bar{w}\}\right\} \cap \bar{N} = \emptyset.$$

Take

$$y \in \left\{\Gamma_{\mathcal{S}}(x) \setminus \{w\}\right\} \cap \left\{\Gamma_{\mathcal{S}}(\bar{x}) \setminus \{\bar{w}\}\right\} \cap \bar{N}.$$

Since $(x, w), (\bar{x}, \bar{w}), (x, y), (\bar{x}, y) \in \mathcal{S}$ and the Kantorovich potentials are differentiable on $\bar{M}$ and $\bar{N}$ we have

$$\tilde{D}_x^- c(x, w) \cap \tilde{D}_x^- c(x, y) \neq \emptyset,$$

$$\tilde{D}_x^- c(\bar{x}, \bar{w}) \cap \tilde{D}_x^- c(\bar{x}, y) \neq \emptyset,$$
$$\tilde{D}_y^- c(x, y) \cap \tilde{D}_y^- c(\bar{x}, y) \neq \emptyset.$$

Therefore, the set $\big\{(x, w), (x, y), (\bar{x}, y), (\bar{x}, \bar{w})\big\}$ forms a $(c, \mathcal{S})$ chain of length 4 containing (x, w) and $(\bar{x}, \bar{w})$. Since $y \in \cup_l N_1$, there exists $l \in \mathbb{N}$ such that $y \in N_l$. Set $l_1 = i(x)$ and $l_2 = i(\bar{x})$. We first assume that $l = 1$. In this case $l = l_1 = l_2 = 1$. Since $(\bar{x}, y)$ is the starting point of the chain

$$\big\{(\bar{x}, y), (\bar{x}, \bar{w})\big\},$$

and $\ell_{\mathcal{S}}(y) = 1$, it means that the above chain is not Right-Maximal and therefore there exists an infinite $(c, \mathcal{S})$-chain

$$S_1 = \big\{(x_1, y_1), (x_2, y_2), ...\big\}$$

such that

$$(x_1, y_1) = (\bar{x}, y), \quad (x_2, y_2) = (\bar{x}, \bar{w}).$$

Similarly, there exists an infinite $(c, \mathcal{S})$-chain

$$S_2 = \big\{(\bar{x}_1, \bar{y}_1), (\bar{x}_2, \bar{y}_2), ...\big\}$$

such that

$$(x_1, y_1) = (x, y), \quad (x_2, y_2) = (x, w).$$

Note that the infinite chains S_1 and S_2 form an infinite $(c, \mathcal{S})$-cycle as the set

$$\big\{(x, w), (x, y), (\bar{x}, y), (\bar{x}, \bar{w})\big\},$$

is a $(c, \mathcal{S})$-chain. This leads to a contradiction.

We now assume $l \geq 2$. It then follows from the construction of N_l that there exits a Right-Maximal $(c, \mathcal{S})$-chain of length l, i.e.,

$$T = \big\{(p_1, q_1), (p_2, q_2), ..., (p_l, q_l)\big\}$$

with $q_1 = y$. By the definition of $f_{\mathcal{S},c,P}$ we must that $l \geq \max\{l_1, l_2\}$. We shall show that at least one of the following assertions hold:

- There exists a Right-Maximal $(c, \mathcal{S})$-chain S with the length bigger than l starting with either (x, y) or $(\bar{x}, y)$.
- There exists a Right-Maximal $(c, \mathcal{S})$-chain S_1 with the length bigger than l_1 starting with (x, w).
- There exists a Right-Maximal $(c, \mathcal{S})$-chain S_2 with the length bigger than l_2 starting with $(\bar{x}, \bar{w})$.

This then leads to a contradiction since $\ell_{\mathcal{S}}(y) = l$, $\ell_{\mathcal{S}}(w) = l_1$ and $\ell_{\mathcal{S}}(\bar{w}) = l_2$. We consider two cases.
Case I. $q_2 \neq y$:

- If $p_1 \neq x$ then consider the Right-Maximal $(c, \mathcal{S})$-chain S as follows:

$$S = \big\{(x, y), (p_1, q_1), ..., (p_l, q_l)\big\}.$$

  and note that $length(S) = l + 1 > l$.

- If $p_1 \neq \bar{x}$ then consider the Right-Maximal $(c, \mathcal{S})$-chain S as follows:

$$S = \big\{(\bar{x}, y), (p_1, q_1), ..., (p_l, q_l)\big\}.$$

Case II. $q_2 = y$:

- If $p_2 \neq x$ then consider the Right-Maximal $(c, \mathcal{S})$-chain S_1 as follows:

$$S_1 = \big\{(x, w), (x, y), (p_2, q_2), ..., (p_l, q_l)\big\}.$$

  and note that $length(S_1) = l + 1 > l_1$.

- If $p_2 \neq \bar{x}$ then consider the Right-Maximal $(c, \mathcal{S})$-chain S_2 as follows:

$$S_2 = \big\{(\bar{x}, \bar{w}), (\bar{x}, y), (p_2, q_2), ..., (p_l, q_l)\big\}.$$

This completes the proof. $\square$.

Remark 6.11. *Note that in the proof of Theorem 6.9 we have only used the generalized twist-condition to ensure that the set*

$$\operatorname{argmax}\big\{c(x, y);\ y \in \Gamma_{\mathcal{S}}(x) \cap N_{i(x)}\big\},$$

is non-empty for all $x \in Dom(f_{\mathcal{S},c})$. Thus one may be able to drop this condition in specific problems such that the non-emptiness of the latter set is otherwise guaranteed.

6.2 Quadratic cost on nested strictly convex sets

Let $\{\mathcal{S}_l\}_l^L$ be a countable (finite or infinite) family of nested strictly convex sets in $\mathbb{R}^{n+1}$. This means that each $\mathcal{S}_l$ is the boundary of a bounded convex body Ω_l and we have the inclusion

$$\Omega_l \subset \Omega_{l+1}.$$

If the sequence is infinite we let $L = \infty$. Note that $\cup_l^L \Omega_l$ is not necessarily a bounded subset of $\mathbb{R}^{n+1}$. The Euclidean norm and the inner product in $\mathbb{R}^{n+1}$ are denoted by $|.|$

and $\langle .,. \rangle$ respectively. Let M be a compact manifold in $\mathbb{R}^{n+1}$, $\mathcal{S} = \cup_l \mathcal{S}_l^L$ and consider the cost function $c : M \times N \to \mathbb{R}$ given by

$$c(x,y) = \frac{1}{2}|x-y|^2.$$

Let μ be a probability measure on X and ν be a probability measure on $\mathcal{S}$. We shall consider the following problem,

$$\inf\Big\{\int_{M\times N} c(x,y)\,d\gamma;\ \gamma\in\Pi(\mu,\nu)\Big\}. \tag{6.2}$$

Our goal is to prove uniqueness and characterize solutions of the above problem. We have the following result for the uniqueness.

Theorem 6.12. *Let $\{\mathcal{S}_l\}_l^L$ be a sequence of nested strictly convex sets in $\mathbb{R}^{n+1}$ with $L \in \mathbb{N}\cup\{+\infty\}$. Let $N = \cup_l \mathcal{S}_l^L$ and assume that the probability measure ν on N is absolutely continuous in each coordinate chart on N and has a finite second moment. Let M be a bounded Borel measurable subset of $\mathbb{R}^{n+1}$ and let μ be a probability measure in $\mathbb{R}^{n+1}$ supported in M. Then any strongly (c,μ,ν)-minimizing set $\mathcal{S}$ is a set of uniqueness.*

We shall need the following result before proving this theorem. At each point $y \in N$, let $n(y)$ be the unit outward normal to N at point y.

Lemma 6.13. *There exists a full ν-measure subset $\bar{N}$ of N such that for all $y \in \bar{N}$ if there exist $x_1, x_2 \in M$ with (x_1,y) and $(x_2,y) \in \mathcal{S}$ then there exists $\alpha \in \mathbb{R}$ such that $n(y) = \alpha(x_1 - x_2)$.*

Proof. Since $\mathcal{S}$ is a strongly (c,μ,ν)-minimizing set, there exist Borel measurable functions $\phi : N \to [-\infty,+\infty)$ and $\psi : M \to (-\infty,+\infty]$ with

$$\psi(x) = \sup_{y\in N}\{\phi(y) - c(x,y)\} \qquad \& \qquad \phi(y) = \inf_{x\in M}\{c(x,y)+\psi(x)\},$$

such that,

$$\mathcal{S} \subseteq \big\{(x,y)\in M\times N;\ \phi(y)-\psi(x) = c(x,y)\big\}.$$

Since M is bounded, it follows from Lemma C.1 in [8] that ψ is locally Lipschitz on N. Let $\bar{N} = Dom(D\phi)$. It follows from the absolute continuity of ν that $\nu(\bar{N}) = 1$. For $y \in \bar{N}$ if there exist $x_1, x_2 \in X$ with (x_1,y) and $(x_2,y) \in \mathcal{S}$, then we must have $D_2c(x_1,y) = D_2c(x_2,y)$ from which the result follows. □

Proof of Theorem 6.12. We shall apply Theorem 5.3. Let $\mathcal{S}$ be strongly (c,μ,ν)-minimizing set. We just need to show that the set $\mathcal{S}$ is $(c,P)-$extreme where $P = \{\mathcal{S}_l\}_l$. As in Definition 2.11, define the maps $i : Dom(f_{\mathcal{S},c}) \to \mathbb{N}$ and $f_{\gamma,c,P} : M \to 2^N$ by

$$i(x) = \min\{l\in\mathbb{N};\ \Gamma_{\mathcal{S}}(x)\cap\mathcal{S}_l \neq \emptyset\}$$

and

$$f_{\gamma,c,P}(x) = \operatorname{argmax}\big\{c(x,y);\, y\in\Gamma_{\mathcal S}(x)\cap \mathcal S_{i(x)}\big\}.$$

Consider the full measure subset $\bar N$ of N given in Lemma 6.13 and the full measure subset $\bar M := M_1$ given in Definition 5.1. Since each $\mathcal S_l$ is compact and c is continuous, condition (i) of Definition 2.11 follows. To prove condition (ii) in Definition 2.11, take distinct points $x_1, x_2 \in \bar M$. Without loss of generality we assume that $i(x_1)\le i(x_2)$. Take $y_1 \in f_{\mathcal S,c,P}(x_1)$ and $y_2\in f_{\mathcal S,c,P}(x_2)$. We need to show that

$$\big\{\Gamma_{\mathcal S}(x_1)\setminus\{y_1\}\big\}\cap\big\{\Gamma_{\mathcal S}(x_2)\setminus\{y_2\}\big\}\cap \bar N=\emptyset,$$

where $\bar N$ is the set given in Lemma 6.13. Take

$$y\in\big\{\Gamma_{\mathcal S}(x_1)\setminus\{y_1\}\big\}\cap\big\{\Gamma_{\mathcal S}(x_2)\setminus\{y_2\}\big\}\cap \bar N.$$

Since $y\in\cup_l\mathcal S_l$, there exists $j\in\mathbb N$ such that $y\in\mathcal S_j$. By the definition of $f_{\mathcal S,c,P}$ we must that $j\ge i(x_2)$. It then follows that $y_1,y_2\in\Omega_j\cup\mathcal S_j$. It also follows from the strict convexity of $\mathcal S_j$ that

$$\langle n(y), y-y_1\rangle>0 \qquad \& \qquad \langle n(y), y-y_2\rangle>0. \tag{6.3}$$

Since $(x_1,y),(x_2,y)\in\mathcal S$, due to Lemma 6.13, there exists $\alpha\in\mathbb R$ such that $n(y)=\alpha(x_2-x_1)$. Substituting this into (6.3) yields that

$$\alpha\langle x_2-x_1, y-y_1\rangle>0, \tag{6.4}$$

and

$$-\alpha\langle x_1-x_2, y-y_2\rangle>0. \tag{6.5}$$

It now follows from the c-cyclical monotonicity of $\mathcal S$ and (6.4) that $\alpha>0$. On the other hand c-cyclical monotonicity of $\mathcal S$ and (6.5) yield that $\alpha<0$ which leads to a contradiction. □

We can also characterize the solutions of the problem (6.2). Let us begin with the following definition.

Definition 6.14. *Say that a measure $\gamma\in\Pi(\mu,\nu)$ is concentrated on the uion of the graphs of measurable maps $\{T_i\}_{i=1}^k$ from M to N, if there exists a sequence of measurable non-negative real functions $\{\alpha_i\}_{i=1}^k$ from M to $\mathbb R$ with $\sum_{i=1}^k\alpha_i(x)=1$ such that for each bounded continuous function $f: M\times N\to\mathbb R$ we have*

$$\int_{M\times N} f(x,y)\,d\gamma=\sum_{i=1}^k\int_M \alpha_i(x) f(x,T_ix)\,d\mu,$$

In this case we write $\gamma=\sum_{i=1}^k(Id\times T_i)_{\#}\mu_i$, where $d\mu_i=\alpha_i\,d\mu$.

Here is our characterization result for the solution of the problem (6.2).

Theorem 6.15. *Let $M \subset \mathbb{R}^{n+1}$ be a compact manifold of dimension n and, μ be non-atomic and absolutely continuous in each coordinate chart on M. Suppose $L \neq \infty$ and ν is absolutely continuous in each coordinate chart on $\mathcal{S} = \cup_l \mathcal{S}_1^L$. Then the problem (6.2) has a unique solution γ and, there exist $k \leq 2L$ and measurable maps, $T_1, ..., T_k : X \to \mathcal{S}$ such that γ is concentrated on $\cup_{i=1}^k Graph(T_i)$.*

We shall need some preliminaries before proving this theorem. We recall the following definition from [16].

Definition 6.16 (**m-twist condition:**). *Let $c : M \times N \to \mathbb{R}$ be a function such that $x \to c(x, y)$ is differentiable for all $y \in N$. Let $m \in \mathbb{N}$. We say that c satisfies the m-twist condition if for each $x_0 \in M$ and $y_0 \in N$ the cardinality of the set*

$$\Big\{y;\ D_x c(x_0, y) = D_x c(x_0, y_0)\Big\},$$

is at most m.

The following characterization of optimal plans of (MK) is established in [16].

Theorem 6.17. *Let M be a complete separable Riemannian manifold and N be a Polish space equipped with Borel probability measures μ on M and ν on N. Let $c : M \times N \to \mathbb{R}$ be a bounded continuous cost function and assume that:*

1. *the cost function c satisfies the m-twist condition;*
2. *μ is non-atomic and any c-concave function on M is differentiable μ-almost surely on its domain.*

Then for each optimal plan γ of (MK), there exist $k \in \{1, ..., m\}$, a sequence $\{\alpha_i\}_{i=1}^k$ of non-negative functions from M to $[0, 1]$, and Borel measurable maps $G_1, ..., G_k$ from M to N such that

$$\gamma = \sum_{i=1}^{k} (Id \times G_i)_{\#}\mu_i, \tag{6.6}$$

where $d\mu_i = \alpha_i \, d\mu$ and $\sum_{i=1}^k \alpha_i(x) = 1$ for μ-almost every $x \in M$. Moreover, if $G_i(x) = G_j(x)$ for some $x \in X$ then $\alpha_i(x) = \alpha_j(x)$.

Proof of Theorem 6.15. The uniqness is already addressed in Theorem 6.12. We show that c satisfies the $2L-$twist condition. Fix $(x_0, y_0) \in M \times N$. If for some $y \neq y_0$,

$$D_x c(x_0, y) = D_x c(x_0, y_0),$$

then there exists $\alpha \in \mathbb{R}$ such that $y - y_0 = \alpha n(x_0)$ where $n(x_0)$ is the unit outward normal to M at x_o. It then follows that $y = y_0 + \alpha n(x_0)$. This argument shows that all the points in the set

$$\Big\{y \in \mathcal{S};\ D_x c(x_0, y) = D_x c(x_0, y_0)\Big\},$$

live on a straight line through x_0 in the direction of $n(x_0)$. On the other hand any stright line can intersect the manifold N in at most $2L$ points. This shows that c satisfies the $2L-$twist condition. It also follows from Lemma C.1 in [8] and absolute continuity of μ that any $c-$concave function is is differentiable μ-almost surely on its domain. Therefore, the result follows from Theorem 6.17. □

A Reminder on universally measurable sets

Let X be a Polish space, that is a complete separable metric space, equipped with its Borel σ-algebra $\mathcal{B}(X)$. A Borel probability measure $\mathfrak{m}$ on X is a probability measure which is defined on all Borels sets in X. Its outer measure $\mathfrak{m}^*$, defined by

$$\mathfrak{m}^*(C) := \inf\Big\{\mathfrak{m}(C) \,|\, C \subset B \in \mathcal{B}\Big\} \qquad \forall C \subset X,$$

extends $\mathfrak{m}$ into a complete measure on the σ-algebra $\mathcal{B}^{\mathfrak{m}}$ generated by $\mathcal{B}$ and the set of all sets of $\mathfrak{m}$-measure zero. A subset U of X is called universally measurable if it is measurable with respect to every complete Borel probability measure on X, that is if for every Borel probability measure $\mathfrak{m}$ on X it is contained in $\mathcal{B}^{\mathfrak{m}}$ as defined above. By construction and using the classical characterization of σ-algebras of the form $\mathcal{B}^{\mathfrak{m}}$ we have

Proposition A.1. *The set $\mathfrak{U}$ of universally measurable sets in X is a σ-algebra containing $\mathcal{B}$. Furthermore, a set $U \subset X$ is universally measurable if and only if for every Borel probability measure $\mathfrak{m}$ on X there exist $V, W \in \mathcal{B}$ such that $V \subset U \subset W$ and $\mathfrak{m}(W \setminus V) = 0$.*

Recall that a set $A \subset X$ is said to be analytic if there is a Polish space Y and a Borel set $B \subset X \times Y$ such that $A = \pi^X(B)$ where $\pi^X : X \times Y \to X$ denotes the projection on the X-variable. The following result can be found in [19, Chapter 4].

Proposition A.2. *Let X, Y be Polish spaces, the following properties hold:*

(i) For every analytic set $A \subset X \times Y$, the set $\pi^X(A)$ is analytic.

(ii) For every Borel set $A \subset X$, every Borel map $f : A \to Y$ and every analytic set $B \subset Y$, the set $f^{-1}(B) \subset X$ is analytic.

(iii) The class of analytic sets in X is closed under countable unions and countable intersections.

(iv) Any analytic set in X is universally measurable in X.

B A description for (c, μ, ν)-minimizing sets

This is an independent section and the results are usually broader than of what needed throughout the paper. The main result in this section is Lemma B.7 which has been used frequently in the proofs of Theorems 2.10 and 5.3. However, we shall need several preliminary results before proving this lemma.
Let (X, Σ) be a measurable space. The universal $\sigma-$field corresponding to Σ is defined by

$$\hat{\Sigma} = \cap_\mu \Sigma_\mu,$$

where μ ranges over all finite measures on Σ and Σ_μ denotes the $\mu-$ completion of Σ. Recall that a set is called Analytic if it is the continuous image of a Borel set in a Polish space. Every Analytic set is universally measurable. Analytic sets are also called Souslin sets.
We now recall a graph-conditioned selection theorem known in the literature as the Yankov-von Neumann-Aumann selection theorem. We state the result and for a proof of it, we refer to Hu-Papageorgiou ([10], p.p. 158-159).

Theorem B.1. *If (X, Σ) is a measurable space, Y is a Souslin space, and $F : X \to 2^Y \setminus \{\emptyset\}$ is graph-measurable (i.e. Graph$(F) \in \Sigma \times \mathcal{B}(Y)$), then there exists a $\hat{\Sigma} - \mathcal{B}(Y)$ selector of F.*

Here Borel σ-algebra of Y is denoted by $\mathcal{B}(Y)$. The following result shows that every $(\hat{\Sigma}, \mathcal{B}(Y))$-measurable map has a $(\Sigma, \mathcal{B}(Y))$-measurable representation with respect to a fixed finite measure on Σ ([6], Corollary 6.7.6).

Proposition B.2. *Let μ be a finite measure on a measurable space (X, Σ), let Y be a Souslin space, and let $F : X \to Y$ be a $(\Sigma_\mu, \mathcal{B}(Y))$-measurable mapping. Then, there exists a mapping $G : X \to Y$ such that $G = F$ μ-a.e. and $G^{-1}(B) \in \Sigma$ for all $B \in B(Y)$.*

The following is a measurable version of the Berge maximum theorem ([17], Theorem 6.3.24).

Theorem B.3. *If (X, Σ) is a measurable space, Y is a Souslin space, $c : X \times Y \to \mathbb{R} \cup \{-\infty, +\infty\}$ is a $\Sigma - \mathcal{B}(Y)$ measurable function, and $F : X \to 2^Y \setminus \{\emptyset\}$ is a graph-measurable multifunction, then*

1. *The function* $x \to m(x) = \sup\{c(x,y);\ y \in F(x)\}$ *is* $\hat{\Sigma} - \mathcal{B}(Y)$ *measurable.*

2. *If for all* $x \in X$*, the set* $S(x) = \{y \in F(x);\ m(x) = c(x,y)\}$ *is nonempty, then* $Graph(S) \in \hat{\Sigma} \times \mathcal{B}(Y)$.

The following result shows the relation between measurability of a function and the measurability of its graph ([6], Lemma 6.7.1).

Proposition B.4. *Let* X *and* Y *be Souslin spaces. Then the graph of any Borel mapping* $f : X \to Y$ *is Borel, hence Souslin, subset of the Souslin space* $X \times Y$*. Conversely, if* $f : X \to Y$ *has a Souslin graph, then* f *is Borel measurable.*

We are now in position to prove the following result which is essential for the proof of Lemma B.7.

Proposition B.5. *Let* (X, Σ) *be a measurable space,* Y *be a Souslin space, and* $F : X \to 2^Y \setminus \{\emptyset\}$ *be a set-valued function with* $Graph(F) \in \Sigma \times \mathcal{B}(Y)$*. Let* $c : X \times Y \to \mathbb{R} \cup \{-\infty, \infty\}$ *be a* $\Sigma - \mathcal{B}(Y)$ *measurable function such that for each* $x \in X$ *the set*

$$\big\{ \operatorname{argmax} c(x,y);\ y \in F(x) \big\},$$

is non empty. Then the following assertions hold:

1. *The set-valued function* $x \to argmax\{c(x,y);\ y \in F(x)\}$ *has a selector that is* $\hat{\Sigma} - \mathcal{B}(Y)$ *measurable.*

2. *The set*

$$C = \big\{x \in X;\ \{\operatorname{argmax} c(x,y);\ y \in F(x)\} \textit{ is not a singlton } \big\},$$

is universally measurable. .

Proof. By Theorem B.3 the function

$$x \in X \to m(x) = \sup\{c(x,y);\ y \in F(x)\},$$

is $\Sigma - \mathcal{B}(Y)$ measurable. Since for each $x \in X$ the set $\{\operatorname{argmax} c(x,y);\ y \in F(x)\}$ is non empty we have that the set

$$S_1(x) := \{y \in F(x);\ m(x) = c(x,y)\},$$

is non-empty and therefore $Graph(S_1) \in \hat{\Sigma} \times \mathcal{B}(Y)$ due to Theorem B.3. It now follows from Theorem B.1 that the multivalued map $S_1 : X \to 2^Y \setminus \{\emptyset\}$ has a selector $S : X \to Y$ that is $\hat{\Sigma} - \mathcal{B}(Y)$ measurable. Let μ be a finite measure on Σ. It follows from Proposition B.2 that there exists a $\Sigma - \mathcal{B}(Y)$ measurable function $S_0 : X \to Y$

such that $S_0 = S$ $\mu-$a.e. Since $Graph(S_0) \in \Sigma \times \mathcal{B}(Y)$ and $Graph(S_1) \in \hat{\Sigma} \times \mathcal{B}(Y)$ we have that $Graph(S_1) \setminus Graph(S_0) \in \hat{\Sigma} \times \mathcal{B}(Y)$. It follows that

$$\pi^1\big(Graph(S_1) \setminus Graph(S_0)\big) \in \hat{\Sigma} \subseteq \Sigma_\mu.$$

We now show that

$$\pi^1\big(Graph(S_1) \setminus Graph(S_0)\big) \subseteq C. \tag{B.1}$$

In fact, for $x \in \pi^1\big(Graph(S_1) \setminus Graph(S_0)\big)$ there exists $y \in Y$ such that

$$(x, y) \in Graph(S_1) \setminus Graph(S_0).$$

It implies that $y \neq S_0(x)$ from which we obtain $x \in C$. We also show that

$$C \setminus \pi^1\big(Graph(S_1) \setminus Graph(S_0)\big) \subseteq \{x \in X;\ S_0(x) \neq S(x)\}. \tag{B.2}$$

To prove this take $\tilde{x} \notin \{x \in X;\ S_0(x) \neq S(x)\}$. We shall show that $\tilde{x} \notin C \setminus \pi^1\big(Graph(S_1) \setminus Graph(S_0)\big)$. It follows that $S_0(\tilde{x}) = S(\tilde{x})$. If $\tilde{x} \notin C$ then we are done. If $\tilde{x} \in C$ then there exists $y \neq S_0(\tilde{x})$ such that $(\tilde{x}, y) \in Graph(S_1)$. This indeed yields that $\tilde{x} \in \pi^1\big(Graph(S_1) \setminus Graph(S_0)\big)$ from which we obtain

$$\tilde{x} \notin C \setminus \pi^1\big(Graph(S_1) \setminus Graph(S_0)\big),$$

as desired. It now follows from (B.1) and (B.2) that

$$\begin{aligned} \pi^1\big(Graph(S_1) \setminus Graph(S_0)\big) &\subseteq C \\ &\subseteq \pi^1\big(Graph(S_1) \setminus Graph(S_0)\big) \cup \{x \in X;\ S_0(x) \neq S(x)\}. \end{aligned}$$

Since $\{x \in X_1;\ S_0(x) \neq S(x)\} \in \Sigma_\mu$ is a μ null set the latter inclusions shows that $C \in \Sigma_\mu$. □

Let $\mathcal{S} \subseteq M \times N$ be a strongly (c, μ, ν)-minimizing set. Recall that the set-valued function $\Gamma_{\mathcal{S}} : M \to 2^N$ corresponding to $\mathcal{S}$ is defined by

$$\Gamma_{\mathcal{S}}(x) = \big\{y \in N;\ (x, y) \in \mathcal{S}\big\},$$

with $Dom(\Gamma_{\mathcal{S}}) = \{x \in M; \Gamma_{\mathcal{S}}(x) \neq \emptyset\}$.

Lemma B.6. *Let $(M, \mathcal{B}(M), \mu)$ and $(N, \mathcal{B}(N), \nu)$ be Polish Borel probability spaces. Let $\mathcal{S} \subseteq M \times N$ be strongly (c, μ, ν)-minimizing set and, assume that $\Gamma_{\mathcal{S}} : M \to 2^N$ is the corresponding set-valued function. Then there exists a set-valued function $F : M \to 2^N \setminus \{\emptyset\}$ such that:*

1. *$Dom(F) = M$ and $F(x) = \Gamma_{\mathcal{S}}(x)$ for $\mu-$a.e. $x \in M$.*

2. *Graph(F) is a Borel subset of $M \times M$.*

3. *For each $\gamma \in \Pi(\mu, \nu)$,*

$$\gamma\big(Graph(F) \setminus \mathcal{S}\big) + \gamma\big(\mathcal{S} \setminus Graph(F)\big) = 0.$$

Proof. Since $\Gamma = \mathcal{S}$ is Borel we have that $Dom(\Gamma_{\mathcal{S}}) = \pi^1(\Gamma) \in \hat{\Sigma}$ where $\pi^1 : M \times N \to M$ is the projection on the first variable. Thus, there exists Borel subsets $A, B \in \Sigma$ with $A \subseteq Dom(\Gamma_{\mathcal{S}}) \subseteq B$ and $\mu(B \setminus A) = 0$. It also follows part 2) of Definition 5.1 that there exists a μ full measure set M_1 of M such that $M_1 \subseteq \pi^1(\Gamma)$. Therefore,

$$1 = \mu(M_1) \leq \mu(B).$$

Thus $\mu(B) = \mu(A) = 1$. Take an arbitrary element $y_0 \in N$ and define $F : M \to 2^N \setminus \{\emptyset\}$ by

$$F(x) = \begin{cases} \Gamma_{\mathcal{S}}(x) & x \in A, \\ y_0 & x \notin A. \end{cases}$$

Note that $Graph(F) = \big(\mathcal{S} \cap (A \times N)\big) \cup \big((M \setminus A) \times \{y_0\}\big)$ is Borel measurable. We now prove assertion (3). It follows that

$$Graph(F) \setminus \mathcal{S} = \Big(\big(\Gamma \cap (A \times Y)\big) \cup \big((X \setminus A) \times \{y_0\}\big)\Big) \setminus \mathcal{S} \subseteq \big((X \setminus A) \times \{y_0\}\big),$$

from which we have

$$\gamma\big(Graph(F) \setminus \mathcal{S}\big) \leq \gamma\big((X \setminus A) \times \{y_0\}\big) \leq \gamma\big((X \setminus A) \times Y\big) = \mu(X \setminus A) = 0.$$

We also have that

$$\mathcal{S} \setminus Graph(F) = \mathcal{S} \setminus \Big(\big(\Gamma \cap (A \times Y)\big) \cup \big((X \setminus A) \times \{y_0\}\big)\Big) \subseteq (X \setminus A) \times Y,$$

and therefore $\gamma\big(\mathcal{S} \setminus Graph(F)\big) = 0$ as desired □

We are now ready to provide a description for strongly (c, μ, ν)-minimizing and c-extreme sets defined in Definitions 5.1 and 5.2 respectively.

Lemma B.7. *Let $c : M \times N \to (-\infty, +\infty]$ be a Borel measurable cost and, let μ, ν be two Borel probability measures respectively on M and N. If $\mathcal{S}$ is strongly (c, μ, ν)-minimizing and strongly c-extreme then there exist Borel measurable functions $h : Dom(h) \subset \bar{M} \to N$ and $g : Dom(g) \subset \bar{N} \to M$ with the following properties:*

(i) There exists a full measure subset $\tilde{M}$ of $\bar{M}$ such that

$$\mathcal{S}' := \mathcal{S} \cap \left(\tilde{M} \times \bar{N}\right) \subset Graph(h) \cup Antigraph(g).$$

(ii) $Graph(h) \cap Antigraph(g) = \emptyset$.

(iii) $Dom(h) \subset D(\mathcal{S}, c)$ and $h = f_{\mathcal{S},c}$ on $Dom(h)$ where

$$D(\mathcal{S}, c) := \Big\{x \in M \mid \{argmax\ c(x,y);\ y \in \Gamma_{\mathcal{S}}(x)\} \ \textit{is a singleton}\ \Big\}.$$

Proof. By assumption $\mathcal{S}$ is strongly c-extreme. Without loss of generality we can assume that $Dom(\Gamma_{\mathcal{S}}) = M$ due to Lemma B.6. By part 1) of Definition 5.2 we have that the set

$$\big\{\operatorname{argmax} c(x,y);\ \ y \in \Gamma_{\mathcal{S}}(x)\big\}$$

is non empty for every $x \in M$. It now follows from Proposition B.5 that the set-valued function

$$x \to \operatorname{argmax} \big\{c(x,y);\ \ y \in \Gamma_{\mathcal{S}}(x)\big\},$$

has a selector $S : M \to N$ that is $\hat{\mathcal{B}}(M) - \mathcal{B}(N)$ measurable where $\hat{\mathcal{B}}(M)$ is the $\sigma-$field of universally measurable sets with respect to $\mathcal{B}(M)$. It also follows from Proposition B.2 that there exists a $\mathcal{B}(M) - \mathcal{B}(N)$ measurable function $S_0 : M \to N$ such that $S_0 = S$ $\mu-$a.e. Set

$$X_S = \{x \in M;\ S(x) = S_0(x)\},$$

and note that $X_S \in \mathcal{B}(M)_\mu$. Thus, there exists a Borel measurable set $\tilde{X}_S \subseteq X_S$ with $\mu(\tilde{X}_S) = 1$. By Proposition B.5, the set

$$C_{\mathcal{S}} := \Big\{x \in M;\ \big\{\operatorname{argmax} c(x,y);\ y \in \Gamma_{\mathcal{S}}(x)\big\} \text{ is not a singlton } \Big\},$$

is universally measurable. Therefore, there exist Borel measurable sets C_0, C_1 with $C_0 \subseteq C_{\mathcal{S}} \subseteq C_1$ such that $\mu(C_1 \setminus C_0) = 0$. Note that $M \setminus C_1 \subseteq D(\mathcal{S}, c) \subseteq M \setminus C_0$. Define the Borel measurable function $h : Dom(h) \subset M \to N$ by $h(x) = S_0(x)$ with

$$Dom(h) = (M \setminus C_1) \cap \tilde{X}_S \cap \bar{M},$$

where $\bar{M} \subset M$ is the full measure subset given in Definition 5.2. Note that $Graph(h)$ is a Borel measurable subset of $M \times N$ as the function $h : Dom(h) \subset M \to N$ is Borel measurable. Set

$$A_{\mathcal{S}} := \Big\{ \bigcup_{x \in Dom(h)} \{\Gamma_{\mathcal{S}}(x) \setminus h(x)\}\Big\} \cup \Big\{ \bigcup_{x \in \bar{M} \cap C_0} \Gamma_{\mathcal{S}}(x)\Big\}.$$

Define the map $g : Dom(g) \subset N \to M$, with $Dom(g) = A_{\mathcal{S}} \cap \bar{N}$, by $g(y) = x$ provided $y \in \Gamma_{\mathcal{S}}(x)$ for some $x \in \bar{M}$.

Claim: g is a function.
To prove the claim we first show that if $g(y) = x$ then there exists $y_1 \in f_{\mathcal{S},c}(x)$ with $y \neq y_1$. To see this we consider two cases $x \in D(\gamma, c)$ and $x \notin D(\gamma, c)$. If $x \in D(\gamma, c)$, it follows from $y \in Dom(g)$, $Dom(h) \subseteq \bar{M} \cap D(\mathcal{S}, c)$ and $\bar{M} \cap C_0 \subseteq \bar{M} \setminus D(\gamma, c)$ that $y \in \Gamma_{\mathcal{S}}(x) \setminus h(x)$ and we are done since $h(x) \in f_{\mathcal{S},c}(x)$. If $y \notin D(\gamma, c)$, it follows from $y \in Dom(g)$ that $x \in \bar{M} \cap C_0 \subset C_{\mathcal{S}}$. Thus, it follows that $\big\{ \operatorname{argmax} c(x,z);\, z \in \Gamma_{\mathcal{S}}(x) \big\}$ is not a singleton. Thus there exists at least one $y_1 \in \big\{ \operatorname{argmax} c(x,z);\, z \in \Gamma_{\mathcal{S}}(x) \big\}$ such that $y \neq y_1$ and keep in mind that $y_1 \in f_{\mathcal{S},c}(x)$. We now show that g is a function. If g is not a function then there exists $y \in Dom(g)$ and distinct points x_1 and x_2 such that $g(y) = x_1$ and $g(y) = x_2$. It follow from the latter argument that there exist $y_1 \in f_{\mathcal{S},c}(x_1)$ and $y_2 \in f_{\mathcal{S},c}(x_2)$ with $y \neq y_1$ and $y \neq y_2$. It then follows that

$$y \in \big\{\Gamma_{\mathcal{S}}(x_1) \setminus \{y_1\}\big\} \cap \big\{\Gamma_{\mathcal{S}}(x_2) \setminus \{y_2\}\big\},$$

which is a contradiction due to condition 2) in Definition 5.2. This completes the proof of the claim.
It is also easily seen that

$$\Big((M \times \bar{N}) \cap Graph(h)\Big) \cup Antigraph(g) = \mathcal{S} \cap (\tilde{M} \times \bar{N}), \tag{B.3}$$

where

$$\tilde{M} := \bar{M} \cap \Big(Dom(h) \cup C_0\Big).$$

We now show that

$$Graph(h) \cap Antigraph(g) = \emptyset. \tag{B.4}$$

Indeed, if $(x, y) \in Graph(h) \cap Antigraph(g)$ then $x \in Dom(h)$ and $y = h(x)$. On the other hand $y \in Dom(g)$ and $g(y) = x$ from which we obtain that $y \in \Gamma_{\mathcal{S}}(x) \setminus \{h(x)\}$ and this leads to a contradiction.

It now follows from (B.3) and (B.4) that

$$Antigraph(g) = \mathcal{S} \cap (\tilde{M} \times \bar{N}) \setminus \Big((M \times \bar{N}) \cap Graph(h)\Big)$$

from which we have that $Antigraph(g)$ is measurable. Therefore, the measurability of the function g follows from Proposition B.4. It now follows from (B.3) that

$$\mathcal{S} \cap (\tilde{M} \times \bar{N}) \subset Graph(h) \cup Antigraph(g).$$

$\square$

C Partial numbered limb systems

A partial numbered limb system is a numbered limb system whose first limb is not necessarily a graph.

Definition C.1 ($\mathcal{S}$-partial numbered limb system). *The set $\mathcal{S} \subset M \times N$ is said to admit a numbered limb system if there are countable disjoint decomposition of M and N,*

$$M = \bigcup_{i=0}^{\infty} I_{2i+1} \quad \textit{and} \quad N = \bigcup_{i=0}^{\infty} I_{2i},$$

with sequences of maps $\{f_{2i+1}\}_{i\in\mathbb{N}^*}$ and $\{f_{2i+2}\}_{i\in\mathbb{N}}$ of the form

$$f_{2i+1} : Dom(f_{2i+1}) \subset I_{2i+1} \subset M \longrightarrow Ran(f_{2i+1}) \subset I_{2i} \subset N$$

and

$$f_{2i+2} : Dom(f_{2i+2}) \subset I_{2i+2} \subset N \longrightarrow Ran(f_{2i+2}) \subset I_{2i+1} \subset M,$$

and a set

$$\mathcal{S}_1 \subset \mathcal{S} \cap (I_1 \times I_0)\,,$$

such that

$$\mathcal{S} = \mathcal{S}_1 \cup \bigcup_{i=1}^{\infty} \left(Antigraph(f_{2i}) \cup Graph(f_{2i+1})\right). \tag{C.1}$$

Each set $\mathcal{S}_1$, $Antigraph(f_{2i})$ and $Graph(f_{2i+1})$ with $i \in \mathbb{N}^*$ is called a limb of $\mathcal{S}$.

Theorem C.2. *Suppose that a Borel set $\mathcal{S}$ admits a partial numbered limb system*

$$\mathcal{S} = \mathcal{S}_1 \cup \bigcup_{i=1}^{\infty} \left(Antigraph(f_{2i}) \cup Graph(f_{2i+1})\right)$$

with the property that $\mathcal{S}_1$, $Antigraph(f_{2i})$ and $Graph(f_{2i+1})$ with $i \in \mathbb{N}^*$ are Borel subsets of $M \times N$. Then all $\gamma \in \Pi(\mu, \nu)$ satisfying $\gamma(\mathcal{S}) = 1$ coincide over $\mathcal{S} \setminus \mathcal{S}_1$.

Proof. Let $\mathcal{S}$ be a set admitting a partial numbered limb system

$$\mathcal{S} = \mathcal{S}_1 \cup \bigcup_{i=1}^{\infty} \left(\text{Antigraph}(f_{2i}) \cup \text{Graph}(f_{2i+1})\right)$$

with $\mathcal{S}_1$, $\{f_{2i+1}\}_{i\in\mathbb{N}^*}$ and $\{f_{2i+2}\}_{i\in\mathbb{N}}$ as in Definition C.1 and such that $\mathcal{S}_1$, Antigraph(f_{2i}) and Graph(f_{2i+1}) with $i \in \mathbb{N}^*$ are Borel subsets of $M \times N$. Set

$$\mathcal{L}_{2i+1} := \text{Graph}(f_{2i+1}) \quad \text{and} \quad \mathcal{L}_{2i} := \text{Antigraph}(f_{2i}) \qquad \forall i \in \mathbb{N}^*.$$

and denote for every $i \in \mathbb{N}^*$ by ν_{2i} the restriction of ν to the set $I_{2i} \subset N$ and by μ_{2i+1} the restriction of μ to $I_{2i+1} \subset M$. Moreover, for every $i \in \mathbb{N}^*$, we denote by $(f_{2i+1})_\sharp \mu$ the push-forward of μ by f_{2i+1}, that is

$$(f_{2i+1})_\sharp \mu := (\mathrm{Id}, f_{2i+1})_\sharp \, \mu$$

which means that for every Borel set $B \subset M \times N$ we have

$$(f_{2i+1})_\sharp \mu(B) := \mu\left(\pi^1(B \cap \mathcal{L}_{2i+1})\right).$$

We note that since $\mathcal{L}_{2i+1}$ is a Borel set, its projection is analytic so it is μ-measurable). In the same way, for every $i \in \mathbb{N}^*$, we denote by $(f_{2i})_\sharp \nu$ the push-forward of ν by f_{2i}. Consider now some plan $\gamma \in \Pi(\mu, \nu)$ such that $\gamma(\mathcal{S}) = 1$ and denote for every integer $k \geq 2$ by γ_k the restriction of γ to $\mathcal{L}_k$. We claim that for every $i \in \mathbb{N}^*$, there holds

$$\left\{ \begin{array}{rcl} \mu_{2i+1} & = & \gamma_{2i+1} + (f_{2i+1})_\sharp \mu_{2i+1} \\ \nu_{2i} & = & \end{array} \right.$$

Assume that there is $\gamma \in \Pi(\mu, \nu)$ such that $\gamma(\mathcal{S}) = 1$. $\square$

References

[1] N. Ahmad, H. K. Kim, R. J. McCann. Optimal transportation, topology and uniqueness. *Bull. Math. Sci.* 1 (2011), no. 1, 13-32.

[2] M. Beiglböck, W. Schachermayer, *Duality for Borel measurable cost functions.* Trans. Amer. Math. Soc. 363, (2011), 4203-4224.

[3] V. Beneš and J. Štěpán. The support of extremal probability measures with given marginals. In *Mathematical Statistics and Probability Theory. Volume A: Theoretical Aspects*, M.L. Puri, P. Révész and W. Wertz, eds. D. Reidel Publishing Co., Dordrecht, 1987, pp 33-41.

[4] Y. Brenier, Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.* 44 (1991) 375-417.

[5] S. Bianchini and L. Caravenna, On the extremality, uniqueness, and optimality of transference plans. *Bull. Inst. Math. Acad. Sin. (N.S.)* 4 (2009) 353-454.

[6] V. I. Bogachev, *Measure theory. Vol. I, II.* Springer-Verlag, Berlin, 2007.

[7] F.H. Clarke, Generalized gradients and applications, *Trans. Amer. Math. Soc.* **205** (1975) 247–262.

[8] W. Gangbo and R. J. McCann. The geometry of optimal transportation. *Acta Math.* **177** (1996) 113–161.

[9] K. Hestir and S. C. Williams. Supports of doubly stochastic measures. *Bernouilli* **1** (1995) 217–243.

[10] S. Hu, N. S. Papageorgiou, *Handbook of Multivalued Analysis.* Volume I: Theory, Kluwer, Dordrecht, The Netherlands, (1997).

[11] A. Katok and B. Hasselblatt. *Introduction to the modern theory of dynamical systems.* Encyclopedia of Mathematics and its Applications, vol. 54. Cambridge University Press, Cambridge, 1995.

[12] V. Levin. Abstract cyclical monotonicity and Monge solutions for the general Monge Kantorovich problem. *Set-Valued Anal.* 7 (1999), 7-32.

[13] R.J. McCann, Existence and uniqueness of monotone measure-preserving maps. *Duke Math. J.* 80 (1995), 309-323.

[14] R. J. McCann and L. Rifford. The intrinsic dynamics of optimal transport. *J. Éc. polytech. Math.*, 3:67-98, 2016.

[15] A. Moameni, *Supports of Extremal doubly stochastic measures.* Canad. Math. Bull. 59 (2016), 381-391.

[16] A. Moameni, *A characterization for solutions of the Monge-Kantorovich mass transport problem.* Math. Ann. 365 (2016), no. 3-4, 1279-1304.

[17] N. S. Papageorgiou, S. Th. Kyritsi-Yiallourou, *Handbook of Applied Analysis.* Advances in Mechanics and Mathematics Volume 19, 2009.

[18] L. Rifford. *Sub-Riemannian Geometry and Optimal transport.* Springerbriefs in Mathematics, 2014.

[19] S.M. Srivastava. *A course on Borel sets.* Graduate Texts in Mathematics, Springer-Verlag, New York, 1998.

[20] C. Villani. *Optimal transport, Old and new.* Grundlehren der Mathematischen Wissenschaften, Springer-Verlag, Berlin, 2009.